

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Biochemical Characterization of the Chp1
Chromodomain Binding to the Nucleosome Core and its Role in
Heterochromatin Formation



Manuel Zocco

aus

Mailand, Italien

2016

Erklärung

Die Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Mario Halic betreut.

Eidesstattliche Versicherung

Die Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 10.11.2015

Manuel Zocco

Dissertation eingereicht am 10.11.2015

1. Gutachter: Herr Prof. Dr. Mario Halic

2. Gutachter: Herr Prof. Dr. Roland Beckmann

Mündliche Prüfung am 01.02.2016

SUMMARY

Eukaryotic genomes are organized inside the cell nucleus in a structured macromolecular DNA-protein polymer named chromatin, formed by single discrete units called Nucleosomes. The packing of the genetic information into chromatin allows the efficient regulation of several nuclear processes, such as gene expression and transcription, DNA replication, cell cycle progression, chromosome segregation and DNA damage repair. Chromatin comes in two flavors: a transcriptionally active, more loosened state, called *euchromatin* and a transcriptionally silent or low expressed, more compact state, called *heterochromatin*. The assembly of silent chromatin or *heterochromatin* is fundamental for the regulation of every nuclear process and it is driven in most Eukaryotes by the deposition and the read-out of the histone H3 lysine 9 methylation (H3K9me) post-translational modification (PTM). H3K9me on the nucleosome is specifically bound by chromatin readers called chromodomains (CD) and this recognition is fundamental for the downstream processes that lead to the formation of heterochromatin and shut down the expression of single genes or entire gene clusters. Despite several studies have been done on different chromodomains binding to H3K9me histone tail peptides, to date there was no structural information on how chromodomains interact with their natural binding partners, the H3K9me₃ Nucleosomes. In a preliminary structural study carried out in our laboratory we solved the cryo-electron microscopy (Cryo-EM) structure of the chromodomain of the fission yeast Chp1 protein (Chp1CD) in complex with an H3K9me nucleosome. The structure showed that the Chp1CD interacts not only with the histone H3 tail but also with the histone globular domains in the Nucleosome core, primarily with histone H3. Mutations in the residues of Chp1CD that form the binding interface with the Nucleosome core (two loops in the β -sheet of the domain) caused a drop of the affinity *in vitro* for the H3K9me Nucleosome, which was independent from the histone H3K9me tail interaction. Cells harboring the same Chp1CD loop mutations were defective in silencing centromeric transcripts and maintain the deposition of the H3K9me mark for heterochromatin formation. This indicated that Chp1CD-nucleosome core interaction is fundamental for heterochromatin formation in fission yeast and opened up to the possibility that chromodomains could read multiple histone PTMs, on both the recruiting histone tail and on the nucleosome core. This study substantially contributes to understand how chromodomains interact with chromatin, how much the nucleosome core interaction is conserved among different CDs and how different chromodomain proteins are regulated at the same loci. Understanding how chromodomain readers recognize nucleosomes is fundamental to uncover the basics of gene silencing and heterochromatin formation.

TABLE OF CONTENTS

SUMMARY	2
TABLE OF CONTENTS	3
INTRODUCTION	5
Heterochromatin: a general introduction	8
Heterochromatin formation in <i>Schizosaccharomyces pombe</i>	14
Heterochromatin at centromeres	16
The RNA interference mediated heterochromatin formation	17
Heterochromatin at the Mating type (MAT) locus	22
Heterochromatin at telomeres	23
Heterochromatin formation main players: HP1 and chromodomain proteins	25
The chromodomain: structure and function in heterochromatin regulation	27
The chromodomain proteins of <i>Schizosaccharomyces pombe</i> in heterochromatin formation	31
Aim of the thesis	35
RESULTS.....	36
<i>In vitro</i> reconstitution of the Chp1CD-H3K9me3 Nucleosome complex.....	36
Chp1CD binding to the H3K9me3 Nucleosome depends on the stability of its interacting loops.....	39
Chp1CD binding to the Nucleosome core is required for heterochromatin formation.....	47
Genomic integration of the LOOP1B/2B mutant causes loss of heterochromatin silencing at centromeres	53
Chp1CD-RNA binding stability is influenced by the Nucleosome core interaction	56
Over-expression of Chp1, Swi6, Chp2 and Clr4 in <i>S. pombe</i> cells destabilizes centromeric heterochromatin	58
<i>In vitro</i> reconstitution of the Swi6-H3K9me3 Nucleosome complex.....	60
<i>In vitro</i> reconstitution of the Chp2-H3K9me3 Nucleosome complex	62
<i>In vitro</i> reconstitution of the Clr4-H3K9me3 Nucleosome complex.....	64
Cross-linking of the Swi6-H3K9me3 Nucleosome complex.....	66
Reconstitution of H3K9me3 di-Nucleosomes	69
DISCUSSION	72
The model of interaction between the chromodomain of Chp1 and the Nucleosome core	78
Conclusions and perspectives	80
MATERIALS AND METHODS	82
MATERIALS	82
METHODS	89

Table of Contents

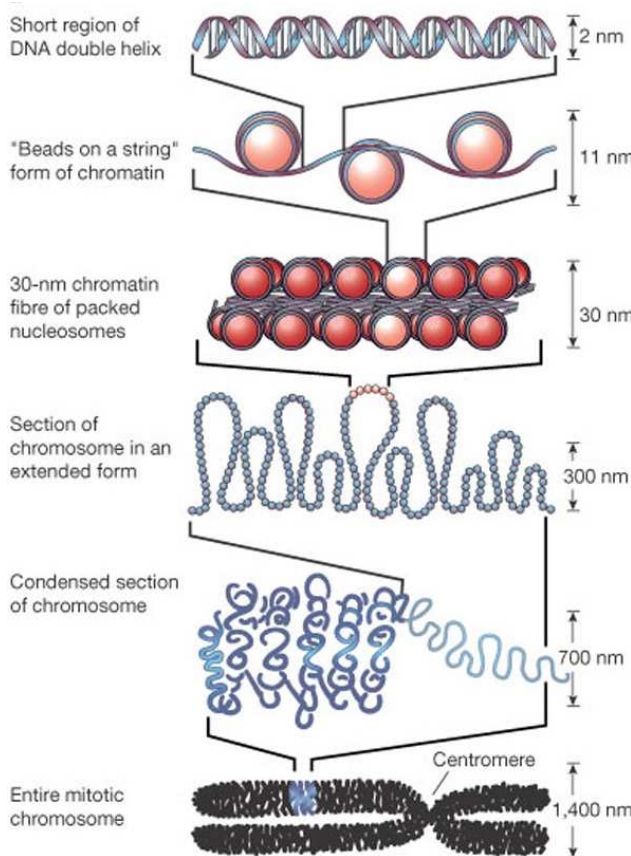
Strains and plasmids construction.....	89
Silencing assays	90
Total RNA purification and Reverse Transcription (RT)	90
Chromatin Immunoprecipitation (ChIP).....	91
Quantitative PCR (qPCR or Real-Time PCR)	92
Protein purification	92
Nucleosome in vitro reconstitution and (H3K ₉ me ₃) methylation.....	93
Chp1CD-H3K ₉ me ₃ MLA Nucleosome in vitro complex formation and elution	94
Nucleosome Trypsin digestion.....	94
Chp1CD-Nucleosome binding assays and Western blot.....	95
Chp1CD-H3K ₉ me ₃ peptide binding assays	96
Microscale Thermophoresis (MST).....	96
RNA electrophoretic mobility shift assays (EMSAs)	97
Negative stain Electron microscopy (EM).....	98
REFERENCES	99
LIST OF ABBREVIATIONS	111
LIST OF FIGURES AND TABLES	112
ACKNOWLEDGEMENTS	115

INTRODUCTION

In order to maintain genome stability and to ensure a finely tuned program of gene expression, Eukaryotes pack their DNA into a higher order macromolecular complex called chromatin. The existence of such organized scaffolds inside the cell nucleus is fundamental to organize the relatively big bulk of genetic material in a functional way, allowing for important processes to occur such as DNA replication, DNA damage repair and transcription.

The basic unit composing chromatin is the nucleosome (Kornberg, 1974; Kornberg and Thomas, 1974; Luger et al., 1997; Olins and Olins, 1974), a DNA – protein complex consisting of four different proteins called histones (H2A, H2B, H3 and H4) forming an octamer (with two copies per each histone), wrapped by 146 bp of DNA (the DNA makes 1.67 left-handed superhelical turns around the core of histone proteins) (Luger et al., 1997) (Figure 1.1). The core of the octamer is formed by a tetramer of histone H3 and H4 (H3-H4)₂ that interact in a four helix bundle mediated by the H3 fold, and from two associated dimers of histone H2A and H2B, that also associate in a four helix bundle with the H3-H4 tetramer at the interface between H2B and H4 (Luger et al., 1997) .

Histones are a family of proteins conserved among Eukaryotes, relatively small (10 to 15 kDa) and highly basic, due to the presence of a substantial number of positively charged residues, particularly lysines and arginines. They are characterized by a globular domain that forms the central core of the nucleosome octamer and a large unstructured part, commonly called histone tail, comprising the first 25-35 residues starting from the N-terminus (Luger et al., 1997). The globular domains of each histone are characterized by the presence of a three helix fold with two loops in between called *histone fold motif*. There are also two structured helical parts named α N in histone H3 and α C in histone H2B that are external to the canonical globular domain, just before the histone unstructured regions (Figure 1.1). Many residues in the histone tail and globular domains are post-translationally modified, and many of these modifications have fundamental roles in regulating gene expression, replication, repair and chromatin organization (Kouzarides, 2007). The high conservation among histone proteins reflects the evolutionary selective pressure in keeping the chemical and structural properties of these proteins that have to act as scaffold for DNA organization in the nucleus and bind many different nuclear factors.



H2B (red), H3 (blue) and H4 (green).

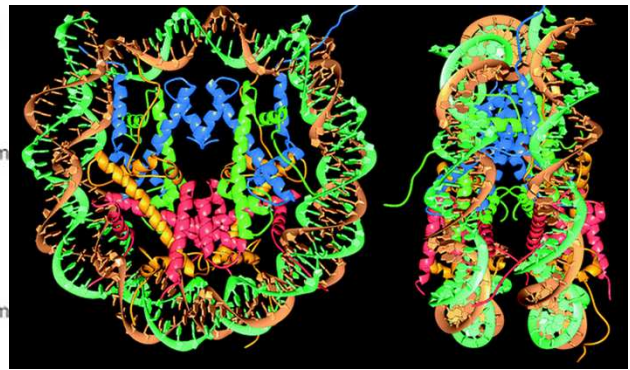


Figure 1.1: The basic unit of chromatin: the Nucleosome core particle

From (Felsenfeld and Groudine, 2003): the model of DNA organization into the chromatin fiber. The DNA is first wrapped around a core of eight histone proteins, forming the Nucleosome particle (shown on the right), the basic unit of chromatin. Nucleosomes are connected with each other through the linker DNA, forming the 11 nm "beads on a string" fiber. Further folding leads to the 30 nm fiber organization, and more complex higher order structures. Chromatin organizations like the 30 nm fiber and higher are still under study with the details of folding still under debate. From (Luger et al., 1997): X-ray Crystal Structure of the Nucleosome core particle at 2.8 Å resolution. 146 bp of DNA are wrapped around a central core of eight histone proteins: H2A (yellow),

The nucleosome forms the main structural component organizing the DNA in the nucleus and each nucleosomal unit is separated by the next with a linker DNA, whose length is varying depending on organisms, tissues and even genome localization (between 20 and 90 bp (Szerlong and Hansen, 2011)). The unfolded structure of the chromatin can be imagined with a "beads on a string" model, with the linker DNA being the "string" connecting each nucleosomal particle with the neighboring (Olins and Olins, 1974, 2003). Within the higher order chromatin organization, another histone protein, histone H1 (and histone H5, only in some organisms), contacts nucleosomes at the DNA entry/exit site and bridges them together, associating in the 30 nm chromatin fiber (Allan et al., 1980; Li and Reinberg, 2011; Lu et al., 2009; Robinson and Rhodes, 2006; Whitlock and Simpson, 1976). Linker histone H1 (and H5) is a large histone protein, with the same properties of the canonical histone forming the nucleosome octamer, bigger in size but much less evolutionarily conserved. Histone H1 has a similar overall histone protein organization, with an N-terminal and C-terminal unstructured domains and a central globular histone fold motif. It associates to the nucleosome, forming the *chromatosome*, to coordinate chromatin organization, and many of the residues at the C-terminal unstructured region of histone H1 have been characterized to be fundamental in

mediating its association and function (Lu et al., 2009).

It is still a matter of debate which nucleosomes are preferentially associated *in vivo*, since different high order chromatin conformations have been isolated *in vitro* with and without histone H1 (Eltsov et al., 2008; Li and Reinberg, 2011; Maeshima et al., 2010; Schalch et al., 2005; Song et al., 2014). A recent cryo-electron microscopy (cryo-EM) study suggests that four nucleosomes are associated together by the linker DNA forming a zigzag twists back and forth among them. These tetrads repetitions are then cross connected by histone H1, which induces a left handed twist while folding the 30 nm fiber, supporting the so-called two start model 30 nm chromatin fiber (Song et al., 2014).

However, recent small angle X-ray scattering (SAXS) studies revealed that no defined 30 nm chromatin fiber exists *in vivo* and most likely the folding observed in previous *in vitro* studies is merely due to the linker DNA length and the 10 nm fiber self-association properties (Maeshima et al., 2014). This however, does not exclude the co-existence of many different types of conformations *in vivo*, including the ones described in the *in vitro* studies mentioned. Each one of the models of 30 nm chromatin fiber proposed might as well be present in the interphase chromatin as their folding depend also on the length of the linker DNA between nucleosomal units (Robinson et al., 2006). The median DNA linker length for the budding yeast *S. cerevisiae* is 23 bp (Brogaard et al., 2012), but it varies across the genome, due to nucleosome positioning variations, nucleosome active remodeling and the binding of transcription factors. In budding yeast the two start fiber model would seem to be the most frequent to occur, but it can be possible to imagine a wide variety of *in vivo* 30 nm fiber conformations depending on nucleosome spacing and local chromatin factors interacting with the DNA, as well as nuclear localization and topological chromosomal architecture. Moreover, histone H1 function in chromatin regulation is not very well characterized in all organisms: for example, the budding yeast *Saccharomyces cerevisiae* seems to require H1 to repress spurious recombination of ribosomal DNA repeats (Li et al., 2008), while it is totally dispensable for cell viability, growth and mating (Ushinsky et al., 1997). On the same hand, the fission yeast *Schizosaccharomyces pombe* is totally lacking H1 and the 30 nm chromatin fiber relies on the folding of the 10 nm fiber for its own stacking (Wood et al., 2002). In yeast, the 30 nm chromatin seems to be able to self-associate independently of histone H1 presence.

To add complexity to the picture, also histone tails, particularly a basic patch of residues on histone H4 tail, seem to be involved in nucleosome arrays compaction (Dorigo et al., 2003). This compaction has been shown to be regulated through acetylation of histone H4K16 residue,

a common post-translational modification occurring *in vivo* (Liu et al., 2011; Shogren-Knaak et al., 2006). Moreover, this single modification is interfering with the activity of the ACF (ATP utilizing Chromatin remodeling and assembly Factor) nucleosome remodeler, even in the presence of ATP, indicating its ability to regulate not only the physical assembly of the 30 nm fiber but also nucleosome positioning (Hwang et al., 2014). Similarly, also histone H3 tail seems to be involved in inter-array nucleosomal interactions, strictly dependent on the state of condensation of the nucleosome arrays and the acetylation status of histone H3 tail (Kan et al., 2007). The position and the spatial organization of the tails is therefore dependent on nucleosome fiber condensation, and post-translational modifications on histone tails might add up an additional level of regulation *in vivo* to determine the higher-order organization of chromatin.

Whether chromatin folds in a defined 30 nm fiber or in an irregular assembly of 10 nm fibers in interphase nuclei it is still a matter of debate (Li and Reinberg, 2011; Maeshima et al., 2010). What is definitely required for the main nuclear processes to occur is a highly dynamic and regulated chromatin conformation with different states of compaction and relaxation.

Chromatin, in fact, comes in two main states, with various intermediates in between: a more compacted state, usually with low level of transcription or no transcription occurring, termed silent chromatin or *heterochromatin*, and a loosened state, genetically active, termed active chromatin or *euchromatin*.

Heterochromatin: a general introduction

In 1928, the German botanist Emil Heitz discovered that chromatin appearance was not the same in interphase nuclei of liverwort cells after staining using a boiling carmin acetate procedure he had devised (Passarge, 1979). In fact he could clearly distinguish more densely stained regions, mainly at the periphery of the nucleus, that he termed *heterochromatin*, in contrast to less stained ones, that he named *euchromatin*.

It was the first cytogenetic evidence of the existence of differential chromatin states. Later, studies on mitotic chromosomes and interphase polytene chromosomes revealed that by staining with particular dyes it was possible to have distinct bands with peculiar patterns correlating with alternating domains of heterochromatin and euchromatin (reviewed in Bickmore and van Steensel, 2013).

On the molecular level there are many features that are distinctive of these two chromatin

states: nucleosome positioning, DNA methylation state, histone post-translational modifications and protein complexes associated. Euchromatin is characterized as being transcriptionally active and replicating throughout the entire S phase, with low levels of DNA methylation, usually localized on chromosome arms and in the inner nuclear regions and recombining during meiosis. There are characteristic histone post-translational modifications associated with actively transcribed regions such as H3K4 methylation, H3K4/K9 acetylation, H3K36 methylation and H3K79 methylation, which are evolutionarily conserved and mark not only single gene expression stages but also global developmental patterns in several model organisms (Kouzarides, 2007) (Figure 1.2).

On the contrary, heterochromatin is usually genetically inert, with no or very low levels of transcription, late replicating in S phase, with no recombination occurring in meiosis, characterized by high levels of DNA methylation and clustered at the nuclear periphery (Figure 1.2). Even gene distribution varies among the two chromatin states, with gene-rich regions being euchromatic and gene-poor regions being often heterochromatic. Many active genes are clustered together facilitating the action of the basal trans-acting transcription factors as well as chromatin remodeling complexes and histone modifying enzymes which are involved in the maintenance of an active gene expression state. In mammalian and bird genomes, gene dense regions are usually A/T poor (Bickmore and van Steensel, 2013; Costantini et al., 2006). Many DNA repetitive sequences are heterochromatic, due to the need of repressing their recombination potential which could eventually lead to genomic instability. It is the case for centromeric repeats in various organisms (like the *dg-dh* repeats in *S.pombe* and the α -satellite repeats in mammalian cells) and various transposable elements. Heterochromatin is characterized by proteins that regulate nucleosomal compaction and gene silencing. The most important proteins belong to the conserved family of the Heterochromatin Protein 1 or HP1 and are the sole proteins structurally characterizing heterochromatin in most organisms (Lomberk et al., 2006a) (Figure 1.2).

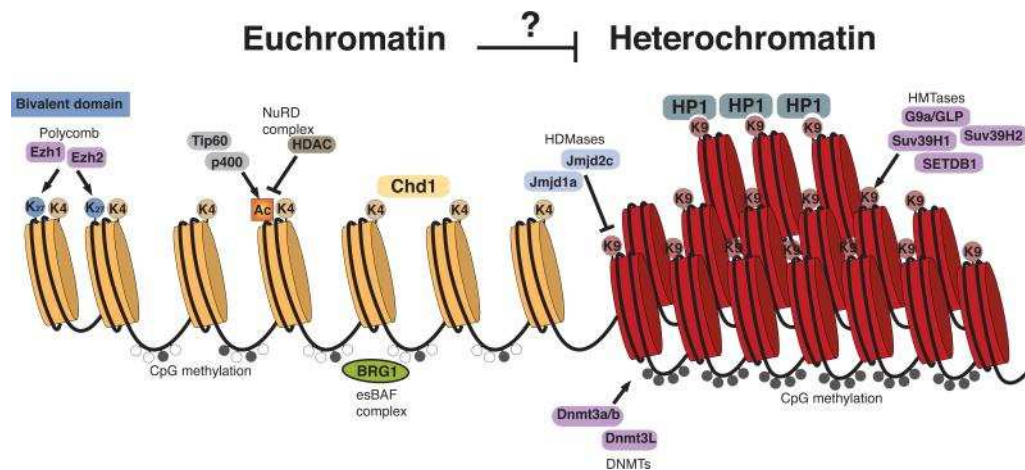


Figure 1.2: Chromatin comes in two main states: Euchromatin and Heterochromatin

From (Gaspar-Maia et al., 2011): overview of the main factors characterizing heterochromatin and euchromatin in mammalian cells. The majority of these protein complexes are conserved from yeast to mammals. Heterochromatin is characterized by a higher degree of compaction, mediated by its structural component, the HP1 proteins. HP1 recognize the H3K9 methylation mark deposited by the Su(Var) 3-9 family SET domain methyltransferases (in the image: Suv39H1, Suv39H2, G9a/GLP, SETDB1, Ctr4 in *S.pombe*). Heterochromatin is also characterized by a high degree of CpG DNA methylation (not conserved in yeast) deposited by the de-novo DNA methyltransferases Dnmt3a/b and Dnmt3L. Boundaries between heterochromatin and euchromatin are set by the removal of the H3K9me mark from the H3K9 de-methylases (in the image: Jmjd1a, Jmjd1c. In *S.pombe*: Epe1). Euchromatin is characterized by a loosened chromatin organization, facilitated by histone acetylation (H3K4ac, H3K9ac, H4K16ac, added by the Tip60, p400 complexes) and the deposition of other active transcription histone marks, such as H3K4 methylation (deposited by the MLL/COMPASS). Some chromatin remodelers, such as the chromodomain – helicase domain Chd1, specifically recognize H3K4 methylation to exert their function. Euchromatin is also characterized by low level of DNA CpG methylation. Intermediate states between Euchromatin and Heterochromatin may exist, in which the simultaneous deposition of counteracting histone marks (such as H3K4 and H3K27 methylation) poises genes in a state of regulated expression that could switch between silencing and active transcription in response to external stimuli. It is the case of the H3K27 methylation, developmentally deposited in facultative heterochromatin by the Polycomb methyltransferases Ezh1 and Ezh2.

The only notable exception is the budding yeast *S. cerevisiae* that does not assemble heterochromatin through HP1 recognition of nucleosomes. HP1 recognize a distinctive histone post-translational mark that is H3K9 methylation, associated with constitutive heterochromatin (Bannister et al., 2001; Nakayama et al., 2001). The recognition happens through specialized protein-protein interaction of the HP1's chromodomain (CD) and the histone H3 lysine 9 methylated residue (Jacobs et al., 2001; Nielsen et al., 2002). Once the recognition is established, HP1 proteins can oligomerize through another domain, the Chromoshadow domain (CSD), thereby self-associating and compacting adjacent nucleosomes (Nishibuchi and Nakayama, 2014; Thiru et al., 2004). Therefore, histone H3K9 methylation is necessary for HP1 dependent heterochromatin assembly. There are other histone post-translational modifications that are heterochromatin specific, such as H3K27 methylation which is deposited by the Polycomb complex PRC2 and is often related to *facultative* or developmentally regulated heterochromatin (Cao et al., 2002; Müller et al., 2002), and global histone H3/H4 deacetylation (like in case of H3K14 and H3K9 deacetylation, necessary for proper centromeric

and MAT locus silencing in *S. pombe* (Ekwall et al., 1997; Grewal et al., 1998)).

In Eukaryotes, proper heterochromatin establishment is crucial to maintain genome stability. At centromeres, heterochromatin is essential to ensure not only the silencing of centromeric repeats, but also to allow the deposition of the centromeric specific variant of histone H3, histone CENP-A, necessary for kinetochore assembly and microtubule attachment both in mitosis and meiosis. In the fission yeast *S. pombe*, heterochromatin is required for CENP-A initial deposition, but not for its maintenance once it is already in place at the centromere (Folco et al., 2008). In fact, deletion mutants of heterochromatin or RNAi factors in *S. pombe* do not coincide with SPB (Spindle Pole Body) attachment defects in mitosis, but have severe SPB attachment and chromosome segregation defects in meiosis, where new centromeres have to re-established, particularly after Prophase I (Klutstein et al., 2015). Constitutive heterochromatin is also found at telomeres, where it is necessary for chromosome ends protection alongside with the Shelterin complex, avoiding chromosome end degradation or chromosome-chromosome fusion (Kano et al., 2005; de Lange, 2009). In *S. pombe*, the Mating type locus (MAT) is constitutively heterochromatic to allow for heterothallic sexual reproduction with a single stable mating-type. The heterochromatin protein Swi6 (the *S. pombe* HP1 protein homolog) function has been initially characterized by the Mating type switching phenotype that occurred in the Swi mutants (Swi stands for Switching), since the MAT locus heterochromatin was completely destabilized and the MAT cassette was de-repressed, causing abnormal meiotic recombination at the MAT type locus (Lorentz et al., 1992).

Constitutive heterochromatin is often located at the periphery of the nucleus, in close contact with the nuclear membrane. Many genes that are silenced and heterochromatic are associated in Lamina Associated Domains (LADs), in close contact with the Lamin proteins of the nuclear membrane lamina (NL) (Padeken and Heun, 2014). Both centromeric and telomeric domains are actually clustered together at the nuclear periphery and growing evidences are suggesting that the association with the nuclear lamina may play a role in heterochromatin formation and maintenance (Bickmore and van Steensel, 2013; Klutstein et al., 2015; Padeken and Heun, 2014). Other heterochromatin regions, rich in H3K9 methylation, are distributed around the nucleolus and are commonly defined as Nucleolus Associated Domains (NADs). It has been hypothesized that heterochromatin distribution between the nucleolus (NADs) and the nuclear envelope (LADs) occurs randomly, and that both the nucleolus and the NL work as platforms for heterochromatin anchoring (Bickmore and van Steensel, 2013; Padeken and Heun, 2014). As already mentioned, DNA hypermethylation, particularly at CpG islands (frequently

occurring at promoters) is a distinctive, but not exclusive, feature of heterochromatin loci. In higher Eukaryotes, DNA methylation at promoters and at the body of genes varies during cell differentiation, developmental stage and in a tissue specific manner and does not always correlate with gene silencing (Suzuki and Bird, 2008). In heterochromatin, once DNA is methylated (usually a 5-methyl-cytosine or 5mC) it can be recognized by specific factors (such as the protein MeCP2 through its methyl-cytosine DNA binding domain – MBD) that can then block transcription by masking the sequences signaling the transcription start, lowering the affinity of transcription factors for their own target sequences or recruiting heterochromatin silencing factors, such as histone deacetylases (Adkins and Georgel, 2011). 5mC deposition is a highly conserved and stable epigenetic mark, found in mammals, plants, *Drosophila* and some yeast species. It is notably absent in both the budding yeast *S. cerevisiae* and the fission yeast *S. pombe*. In mammalian stem cells, where the genomic landscape is strongly euchromatic, DNA is generally hypo-methylated and a similar situation is often seen in several types of cancer, where cells undergo a vast de-differentiation reprogramming in which DNA methylation is partially lost, also at repetitive sequences, giving rise to spurious recombination and genomic instability (Apostolou and Hochedlinger, 2013).

In the plant model *Arabidopsis thaliana*, DNA methylation is coupled with histone H3K9 methylation, in one of the best described heterochromatin formation pathways (Holoch and Moazed, 2015a). Similarly to what happens in the yeast *Neurospora crassa* (Tamaru and Selker, 2001), in *A. thaliana* deposition of H3K9 methylation by the KYP Su(var)3-9 methyltransferase is necessary to recruit the DNA methyltransferase CMT3 to deposit the 5mC onto the DNA (Johnson et al., 2002). Like HP1 proteins, CMT3 is able to recognize specifically H3K9 methylation through its chromodomain and then modify the underlying DNA sequence (Du et al., 2012). In an analogous way, in human primary somatic and germline cells, H3K4 dimethylation was found at hypo-methylated inactive CpG island promoters and it was hypothesized that the mark could act to prevent DNA methylation (Weber et al., 2007). These evidences indicate the presence of a hierarchy between epigenetic marks, with histone modifications being upstream of DNA methylation.

The first genetic evidences for heterochromatin control of gene expression was the characterization of the Position Effect Variegation (PEV). The first experiments performed by Muller in 1930 were showing that by irradiating *Drosophila melanogaster* flies with X-rays it was possible to mutate a gene responsible for eye pigmentation from a red to white phenotype, and that some of the mutant flies eyes were showing a distinct variegated pattern (or mottled), with some facets of the eye red and some white. In a subsequent irradiation step, it was possible

to revert the phenotype back. Since it was not likely that the gene responsible for the eye color had been mutated again, some other mutation occurring in a different gene was responsible for reverting the phenotype (Muller, 1930). It was later discovered that the initial mutation occurred because of an X-ray induced translocation that was placing the *white* gene close to a heterochromatic region, thereby seldom repressing its transcription. Subsequent mutagenesis steps could recover the *white* gene expression by impairing genes directly involved in heterochromatin formation and spreading, such as the suppressor of variegation *Su(var)205* gene encoding for the heterochromatin protein HP1a (Eissenberg et al., 1992).

The evidences provided by these initial experiments indicated that the silencing could be spread across neighboring loci and that the heterochromatin formation was under the control of specific genes. In particular, mutations on Suppressor of variegation *Su(var)* genes, prevents the silencing of a reporter gene nearby a heterochromatic region. Under these category of genes a wide variety of different functions can be found, such as the *Su(var)3-9*, a histone H3K9 methyltransferase (homolog of the *S. pombe* Clr4), and several HP1 proteins.

On the other hand mutation on Enhancers of variegation *E(var)* genes, increase the frequency of silencing of a reporter gene close to a heterochromatin region, indicating a spread beyond the chromatin boundaries as in the case of an H3K9 specific histone demethylase, an enzyme involved in preventing the spread of the heterochromatin H3K9me mark on transcriptionally active regions, like the *S.pombe* Epe1 (Braun et al., 2011; Ragunathan et al., 2014). As can be deduced, heterochromatin boundaries have to be tightly controlled and there are several factors contributing to it. The presence of an actively transcribed gene, for instance, and all histone marks regulating its transcription, such as histone acetylation and H3K4me, are key players in preventing heterochromatin spreading. Moreover, the presence of NFR (Nucleosome Free Regions) impairs the deposition of new H3K9me and other repressive marks as well as a high nucleosome turn-over rate and chromatin remodeling (reviewed in (Wang et al., 2014). The process of organizing heterochromatin involves three main steps: the initial nucleation, the heterochromatin spreading from the initiation sites and the maintenance through the cell divisions. In different organisms, heterochromatin establishment involves different factors that allow the initial deposition of the histone H3K9me (or H3K27me) or the deacetylation of H4K16 (as occurs in *S. cerevisiae*). However, the initial signal that codes for heterochromatin formation passes through a histone post-translational modification deposition (in case of a repressive mark) or removal (in case of histone acetylation). The subsequent spreading occurs through amplification of the initial signal and several cycles of mark deposition or removal across the heterochromatic locus, with specific protein complexes involved in the process, such

as HMT –histone methyltransferases and HDAC- histone deacetylases. A boundary is set when an antagonizing histone post-translational mark is deposited on a nearby region, thus confining the heterochromatin formation at specific loci (Figure 1.2). The carefully controlled balance between nucleosome remodeling, histone PTMs deposition, nucleosome turn-over and transcription is what controls silencing and expression of single protein coding genes. This is therefore fundamental to allow timely and controlled expression of proteins and to prevent the rise of disease conditions, potentially fatal for the cell (Bhaumik et al., 2007; Garcia et al., 2015).

In the fission yeast *S. pombe*, the initial deposition of the H3K9me heterochromatin mark at centromeres is RNA interference (RNAi) dependent. RNAi factors recruit, in fact, a specific HMT (the Cul4-Rik1-Clr4 or CLRC complex) that deposits the H3K9me mark across the whole heterochromatin locus. The H3K9me mark is then recognized and bound by HP1 proteins that cluster together nearby nucleosomes and silence the whole region (Bannister et al., 2001; Lachner et al., 2001; Nakayama et al., 2001). The H3K9me mark deposition and its readout are the key feature for heterochromatin formation, spreading and inheritance in *S.pombe* and in all the organisms, including mammals, in which similar mechanisms have been described. Since RNAi factors, H3K9me, HP1 and chromodomain proteins are well conserved from fission yeast to human, *S.pombe* represents one of the best organisms to study heterochromatin and silencing, given that its genome has been completely mapped, it grows rapidly under laboratory conditions and it is easy to genetically manipulate.

Heterochromatin formation in *Schizosaccharomyces pombe*

As described earlier, there are three main constitutive heterochromatin loci in *S. pombe*: centromeres, telomeres and the Mating type (MAT) locus (Allshire et al., 1995) (Figure 1.3). All three loci share homology between distinct non-coding DNA sequence repeats, named *cenH* like repeats or *cenH* elements, which have a central role in heterochromatin formation. These sequence elements are transcribed at low levels by RNA polymerase II and subsequently targeted by RNAi to induce transcriptional gene silencing (TGS) and H3K9me mark deposition (Djupedal et al., 2005; Noma et al., 2004; Verdel et al., 2004; Volpe et al., 2002).

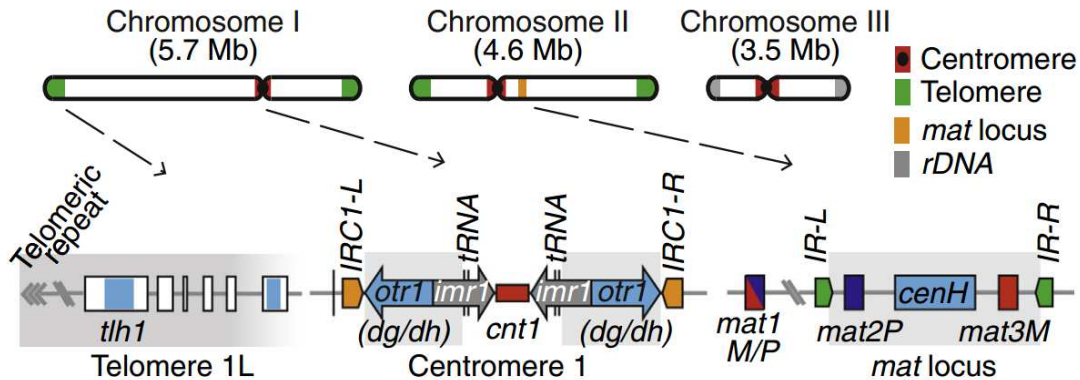


Figure 1.3: Overview of the three main constitutive heterochromatin loci in *S. pombe*

From (Mizuguchi et al., 2015): all three chromosomes in fission yeast have extensive heterochromatin domains that are constitutively maintained throughout cell divisions: the telomeric ends, the centromeres and the Mating-type (*mat*) locus on chromosome II. At centromeres, two classes of repeats, *imr* (innermost) and *otr* (outermost), surround the central core, *cnt*, where the kinetochore is assembled. *Otr* repeats can be further subdivided in *dg* and *dh* repeats, both transcribed and targeted by RNAi for heterochromatin formation. *tRNA* genes and *IRC* inverted repeats form boundary elements, limiting the spreading of heterochromatin from the centromere. At telomeres, multiple pathways are involved in heterochromatin formation, *Taz1* and *Swi6* dependent. RNAi plays a role in the initial establishment of heterochromatin from the sub-telomeric region, containing conserved *dh*-like repeats. Similarly a *cenH* element, homologous to *dg* and *dh* centromeric repeats is present at the *mat* locus, in between the *IR-L* and *IR-R* inverted repeats and the silenced *mat2P* and *mat3M* cassettes. In the blue boxes, the *dg-dh* repeats and their homologous sequence at the three heterochromatin loci.

This process is fundamental in the establishment of heterochromatin formation. Heterochromatin is also found at rDNA repeats (Figure 1.3), even though its role has still not been elucidated in detail (Cam et al., 2005). In addition to constitutive heterochromatin, facultative heterochromatin can be assembled at other loci as well, targeting meiotic genes and other genomic locations in a euchromatic context (Hiriart et al., 2012; Zofall et al., 2012). All these heterochromatin loci share the presence of high levels of H3K9 dimethylation, low level of histone acetylation (H3K4ac, H3K9ac, H3K14ac and H4K16ac) and binding of the two HP1 proteins, Chp2 and Swi6, to the H3K9me2 histone mark. Moreover, reporter genes inserted in the vicinity or within any of these constitutive heterochromatin loci are transcriptionally silenced (Allshire et al., 1995). The H3K9 methylation mark deposition is carried out by the HMT Clr4, the only enzyme specific for the H3K9 residue, generating all three methylation states: mono-, di- and trimethylation (Al-Sady et al., 2013; Nakayama et al., 2001; Rea et al., 2000). Clr4 is able to bind to the H3K9me mark through its chromodomain, thereby coupling the deposition of the H3K9me mark with its read-out, allowing the spreading of heterochromatin to nucleosomes directly nearby and maintain the H3K9me mark at constitutive loci upon cell division and mitotic growth throughout several generations (Al-Sady et al., 2013; Audergon et al., 2015; Nakayama et al., 2001; Ragunathan et al., 2014). While at centromeres RNAi is necessary for heterochromatin establishment and maintenance, at telomeres and at the

MAT locus other pathways are also involved as concurrent mechanisms (Jia, 2004; Kanoh et al., 2005).

Heterochromatin at centromeres

The three centromeres in *S. pombe* have variable size, ranging from 35 to 110 kb (Partridge et al., 2000; Steiner et al., 1993; Volpe et al., 2002; Wood et al., 2002). They are characterized by having a central core, called *Cnt*, the point of kinetochore assembly. In fact, the main function of centromere is ensuring an even chromosome segregation upon cell division in both mitosis and meiosis. To do so, a complex protein machinery called kinetochore has to assemble on the *cnt/cc* central region and then attach to the microtubules of the spindle fiber. The first component of the inner layer of the kinetochore is the histone H3 variant, Cnp1 (CENP-A homolog), that is found specifically in the *cnt/cc* region and additionally in the most internal part of the *imr* (innermost) repeats. The deposition of Cnp1 is dependent on heterochromatin only in the establishment of new kinetochores (Folco et al., 2008; Klutstein et al., 2015). Moreover, the central DNA sequence (around 2 kb- AT rich) in the *cnt/cc* is required for Cnp1 deposition, and the process require specific stalling of RNA Pol II transcription over the centromeric locus. The centromeric sequence itself specifies for the low efficiency transcription and RNA Pol II stalling, which is required for CENP-A Cnp1 deposition (Catania et al., 2015), highlighting that centromere establishment is directed by both genetic and epigenetic mechanisms. Heterochromatin at centromeres has been shown to be facilitating the recruitment of the Rad21-cohesin to ensure sister centromere attachment and proper segregation. The deletion of Swi6 impairs the Rad21-cohesin recruitment causing segregation defects (Bernard et al., 2001). A similar role for Swi6-Psc3 cohesin has been shown at the MAT locus (Nonaka et al., 2002).

Moreover, in a recent study, pericentromeric heterochromatin was found to be essential in recruiting cohesins to promote chromatin fiber compaction and inter-arm interactions within centromere proximal regions, forcing chromosomes in a rigid organization that limits extensive contacts between various chromatin domains. Loss of heterochromatin has the effect of increasing the inter- and intra- chromosomal arm associations, possibly enhancing the likelihood of aberrant recombination or impairing other nuclear processes (Mizuguchi et al., 2014).

The central core *cnt/cc* of the centromere is surrounded by two distinct classes of repeats: the closest to the kinetochore attachment site, the *imr* or innermost repeats, and the *otr*, or

outermost repeats (Partridge et al., 2000; Steiner et al., 1993; Volpe et al., 2002; Wood et al., 2002). Both the *imr* and *otr* repeats are heterochromatic and any reporter gene inserted into these loci gets silenced (Allshire et al., 1995; Partridge et al., 2000). The external boundaries of centromeric heterochromatin are defined by the presence of actively transcribed tRNA genes (5 genes flanking three centromeric loci) (Noma et al., 2006; Scott et al., 2006), and localize anti-silencing factors such as the Epe1 H3K9 demethylase, the bromodomain protein Bdf1 and the histone acetyltransferase Mst2 (Cam and Grewal, 2004; Partridge et al., 2000; Wang et al., 2013b, 2015). tRNA genes are also found in the inner core of the centromere (in chromosome 3) and at the *imr* repeats (chromosomes 1-2-3), but are transcriptionally silenced (Volpe et al., 2002). Functionally, the centromere in *S.pombe* can be divided into two main regions: 1) an inner core composed by the *cnt/cc* kinetochore attachment site and the *imr* repeats and 2) the outer repeats (*otr*) where pericentromeric heterochromatin is established. *Otr* repeats can be further sub-divided into two classes: *dg* and *dh* repeats. These blocks of sequences are present in multiple copies around each centromere and their copy number determines the variable size (35-110 kb) of the three centromeric loci in *S.pombe* (Partridge et al., 2000; Steiner et al., 1993; Volpe et al., 2002; Wood et al., 2002) (Figure 1.3).

All centromeric locus is subject to low levels of transcriptional activity (Catania et al., 2015; Djupedal et al., 2005), which is essential for establishment of pericentromeric heterochromatin in a RNA interference mediated process (Kato, 2005; Verdel et al., 2004; Volpe et al., 2002).

The RNA interference mediated heterochromatin formation

Centromeric transcription generates both sense and antisense transcripts that are targeted by RNA interference in a co-transcriptional gene silencing process (Halic and Moazed, 2010; Kato, 2005; Moazed, 2009; Verdel et al., 2004; Volpe et al., 2002). The first evidences connecting RNAi to centromeric silencing came from the work of the Martienssen laboratory and others, showing that deletion of Argonaute (*ago1*+), Dicer (*dcr1*+) or the RNA dependent RNA polymerase (*rdp1*+) caused upregulation of centromeric transcripts and loss of H3K9 methylation at the centromeric locus (Provost et al., 2002; Volpe et al., 2003, 2002). Further prove of RNA interference involvement in targeting centromeric transcripts was the isolation and characterization of siRNAs (short interfering RNAs) mapping to the pericentromeric *dg* and *dh* repeats (Reinhart and Bartel, 2002). The *otr* pericentromeric repeats are the initial nucleation site for centromeric heterochromatin and the process is dependent on the targeting of centromeric transcripts by RNA interference, directed by the Argonaute complex RITS

(RNAi- induced transcriptional silencing) (Bühler et al., 2006; Verdel et al., 2004). The RITS complex is composed by three proteins: Argonaute (Ago1), binding short interfering RNAs (siRNAs) and targeting centromeric transcripts, the GW (glycine tryptophan) domain protein Tas3 and the chromodomain protein Chp1. The purification of Chp1 gave the first connection between the siRNA targeting of centromeric transcripts and heterochromatin (Verdel et al., 2004). Chp1, in fact, binds H3K9me nucleosomes by direct recognition of the H3K9me mark through its chromodomain, and this interaction is fundamental for heterochromatin establishment and maintenance (Partridge et al., 2000; Schalch et al., 2009; Verdel et al., 2004). The specific targeting of pericentromeric transcripts occurs through base pair complementarity between an Argonaute bound siRNA and its target (Verdel et al., 2004). Argonaute is loaded with double stranded dsRNA in the ARC complex (Argonaute siRNA chaperone), containing Ago1 and two additional proteins, named Arb1 and Arb2, required as chaperones to facilitate the process (Buker et al., 2007; Holoch and Moazed, 2015b). The loaded Ago1 can be then incorporated into the RITS complex. The strand that is selected to be the targeting “guide” strand of the loaded dsRNA is the one more thermodynamically stable of the two, binding the MID domain with the 5′ end (Peek and Behlke, 2007). The other strand or “passenger” strand is usually cut, thanks to Ago1 slicer endonucleolytic activity, removed and targeted for degradation. An Argonaute slicer activity mutant fail to target pericentromeric transcripts, despite loading dsRNA, and has defects in heterochromatin formation (Buker et al., 2007; Halic and Moazed, 2010). The dsRNAs loaded by the ARC chaperone complex are in average double stranded 21-25 nt long with 2 nt overhangs at the 3′ end, bound by the Ago1 PAZ domain, and a 5′ phosphate, bound by the Ago1 MID domain. These short dsRNA are the products of the degradation of self-associated sense and antisense dsRNA pericentromeric transcripts by the type III RNA endonuclease Dicer (Dcr1) (Holoch and Moazed, 2015a; Moazed, 2009; Provost et al., 2002). Once sense and antisense centromeric transcripts are generated they can pair together to generate dsRNA that is then targeted by Dcr1. The product are short dsRNA that can then be loaded into Ago1 through the ARC complex, which is not associating with chromatin, but serves as a platform to generate fully loaded Argonaute that can be then incorporated into the RITS complex (Holoch and Moazed, 2015b). RITS association through its chromodomain protein Chp1 to H3K9me chromatin, while targeting pericentromeric RNAs with Ago1 associated siRNAs, is fundamental to recruit two distinct machineries: one is the RNA dependent RNA polymerase complex or RDRC and the other is the Clr4-Rik1-Cul4 (CLRC) complex.

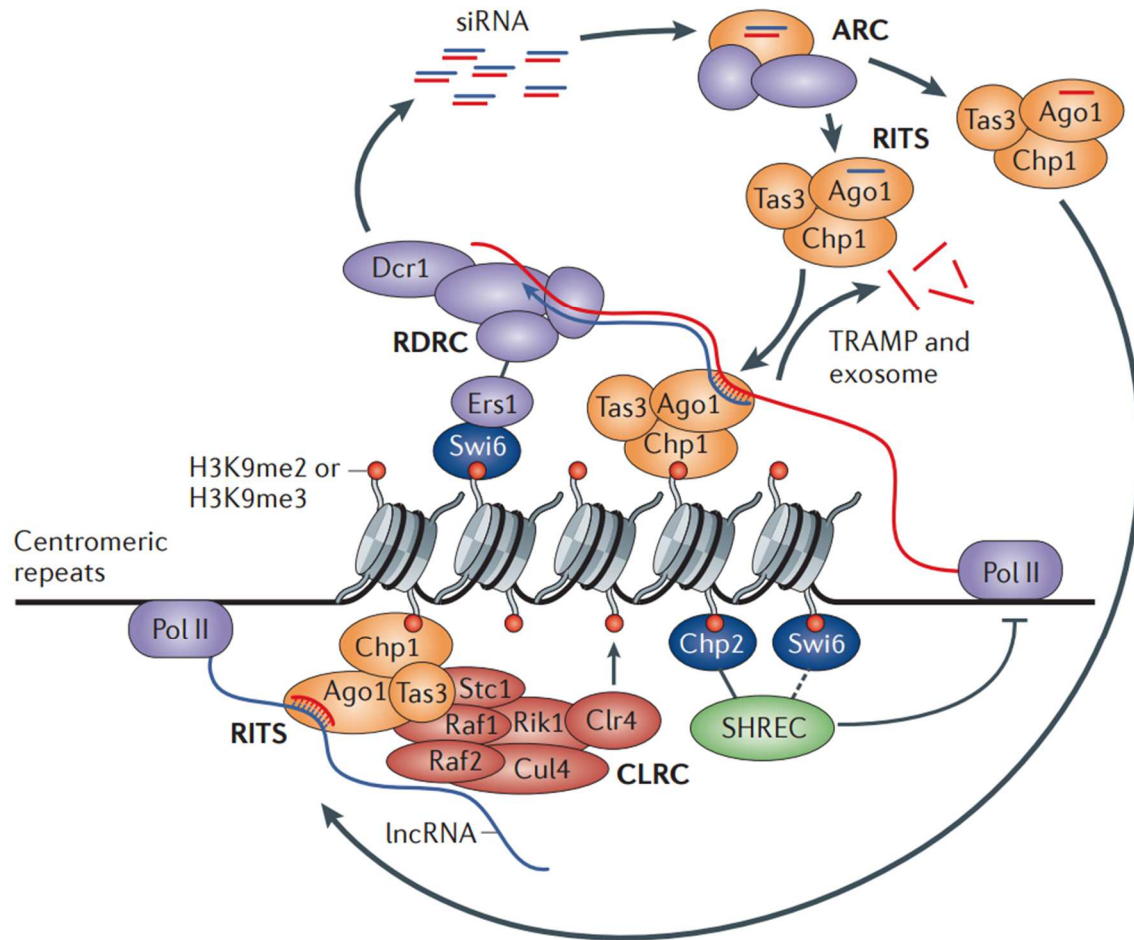


Figure 1.4: The “nascent transcript” positive feedback loop model for heterochromatin formation at centromeres in *S.pombe*

From (Holoch and Moazed, 2015a): Centromeric transcripts (*lncRNA*, transcribed by RNA PolIII) are targeted by the RITS-Argonaute complex (formed by Chp1, Tas3 and Ago1) by complementary siRNAs loaded onto Argonaute (Ago1). The Chp1 protein, associates the Ago1 targeting the centromeric transcripts with the chromatin, though its binding to H3K9 methylated Nucleosomes (H3K9me2 or H3K9me3). Another component of the RITS, the GW protein Tas3, bridges Chp1 and Ago1 together and forms together with Chp1 and the transcribed centromeric *lncRNA* the recruiting platform for the siRNA loaded Ago1. The RITS once targeting the RNA, recruits two complexes: the RDRC and the CLRC. The RDRC (RNA dependent RNA polymerase complex) synthesizes a dsRNA starting from the targeted centromeric RNA and the Ago-bound siRNA as templates. dsRNAs can be then sliced by Dicer (Dcr1) to generate new siRNAs for centromeric RNA targeting, thus creating positive feedback loop. RDRC can be also recruited through Ers1 interaction with the HP1 protein Swi6. The second complex recruited by RITS, through Tas3-Chp1 interaction with Rik1 and Stc1, is the CLRC (Clr4-Rik1-Cul4 complex), containing the SET domain methyltransferase Clr4. Clr4 deposits the H3K9 methylation, which is then bound by the chromodomain of Chp1, and of the HP1 proteins, Swi6 and Chp2, for nucleosome compaction and heterochromatin formation. Both HP1s, Swi6 and Chp2, can recruit an additional silencing complex, the histone deacetylase SHREC complex (Snf2-histone deacetylase repressor complex), specific for H3K14 deacetylation. The TRAMP non-canonical poly(A) polymerase and the exosome contribute as well to the silencing of centromeric transcripts by direct degradation. The binding of the RITS complex to H3K9me Nucleosomes with Chp1 and to the centromeric RNA through Ago1 loaded siRNA, constitutes the core event for heterochromatin formation in the nascent transcript feedback loop model.

Introduction

The RDRC complex is composed by Rdp1, the RNA dependent RNA polymerase, the Hrr1 helicase and the Cid12 non-canonical poly-A polymerase (Motamedi et al., 2004; Sugiyama et al., 2005). The RNA dependent RNA polymerase complex associate with the RITS and catalyzes the synthesis of the complementary strand of the pericentromeric transcripts targeted by Ago1. This generates more substrates for the Dcr1 dependent 21-25nt dsRNA production (Colmenares et al., 2007; Motamedi et al., 2004), thus amplifying the initial input siRNA in a feedback positive loop. A novel protein, Dsh1, recently characterized, helps the activity of Dcr1 in generating siRNA and recruits RDRC complex to perinuclear foci to generate siRNA generating factories (Kawakami et al., 2012). The CLRC complex, on the other hand, interacting with the RITS through the subunit Rik1 (Gerace et al., 2010) and the bridging LIM domain protein Stc1 (Bayne et al., 2010), is involved in depositing the H3K9 methylation mark through its methyltransferase subunit Clr4 (Zhang et al., 2008). Clr4 is a H3K9 specific methyltransferase, able to deposit the H3K9me mark and bind to it through its chromodomain, enabling a dual function of both “writer” and “reader” of the histone modification (Al-Sady et al., 2013; Audergon et al., 2015; Nakayama et al., 2001; Ragunathan et al., 2014; Zhang et al., 2008). The other major components of the CLRC complex are the proteins Rik1 (that associates with both the RDRC and the RITS complex), Raf1, Raf2 and the E3 ubiquitin ligase Cul4. Cul4 target for ubiquitination is still unknown but its activity is required for heterochromatin formation (Jia et al., 2005).

The deposition of the H3K9 methylation, as previously described, creates a binding platform for the other *S. pombe* chromodomain protein: Chp1 and the two HP1 proteins, Swi6 and Chp2. HP1 protein binding contributes to centromeric silencing by bridging together nearby nucleosomes through recognition of the H3K9me mark and self-oligomerization through the chromo-shadow domain, and by recruiting HDAC complexes such as the SHREC and Sir2 deacetylases (Buscaino et al., 2013; Motamedi et al., 2008; Sugiyama et al., 2007). Histone deacetylation contributes to erasing the acetylation mark to probe chromatin for H3K9me deposition. The SHREC complex, formed by the Clr1, Clr2, Clr3 and Mit1 proteins, deacetylates specifically the histone H3K14 residue (through the activity of the H3K14 deacetylase Clr3), and is required for dictating proper nucleosome positioning in heterochromatin and silence the transcriptional activity of RNA Polymerase II (Motamedi et al., 2008; Sugiyama et al., 2007; Wirén et al., 2005). Sir2 specifically de-acetylates the histone H3K9 residue, making it accessible for methylation mark deposition (Wirén et al., 2005). It has been proposed that H3K4 acetylation mediated by the acetyltransferase Mst1, might work as a switch mark to recruit HP1 (Swi6/Chp2) proteins to H3K9me chromatin during

heterochromatin reassembly (Xhemalce and Kouzarides, 2010). Recruitment of the SHREC effector complex by HP1 (Swi6) proteins to H3K9me chromatin is also regulated by post-translational modifications at HP1 N-terminus by the Casein-kinase 2 (CK2). Loss of CK2 dependent phosphorylation at the N-terminus of Swi6 reduces the localization of SHREC to heterochromatin with a concomitant increase of the demethylase Epe1 localization and impairment of the transcriptional gene silencing (TGS) (Shimada et al., 2009). A recent study highlights the evolutionary importance of CK2 phosphorylation to HP1 binding to nucleosome, which is increasing the specificity for the H3K9me mark, by impairing non-specific binding to nucleosomal DNA (Nishibuchi et al., 2014). Chromatin bound Swi6 can also recruit additional RNAi complexes (RITS, RDRC) and Dcr1 through the interaction with the silencing factor Ers1, linking directly heterochromatin to RNA interference (Hayashi et al., 2012; Rougemaille et al., 2008, 2012). Ers1 strongly interacts with the RDRC subunit Hrr1, and loss of Ers1 impairs RDRC centromeric localization. This process of recruitment of the RDRC complex is RNA interference independent and explains the importance of having the Swi6 heterochromatin platform properly assembled to anchor the RDRC complexes to chromatin (Hayashi et al., 2012). Moreover, HP1 protein help recruiting the HIRA/Asf1 histone chaperone complex, which, in concert with the histone deacetylase Clr6 and the SHREC complex, suppresses antisense transcription by promoting histone de-acetylation and promoting nucleosome occupancy at the heterochromatin repeats. TFIIC factor (RNA polymerase III specific initiation factor – tRNA gene specific) act as a boundary element, limiting the spreading of SHREC and Clr6 de-acetylase activity to tRNA bordering genes (Yamane et al., 2011).

One of the biggest questions in the latest years was to find out what is the triggering factor for this positive feedback loop mechanism that establish heterochromatin in *S. pombe*. Halic and Moazed discovered a new class of small RNA, called primal RNAs (priRNAs), which are generated independently of Dicer activity and are able to establish small level of H3K9 methylation by associating with the RNAi machinery. priRNAs are product of general sense and antisense transcript degradation and stochastically associate with Argonaute to initiate RNAi. Because of naturally occurring high level of antisense transcription at heterochromatic loci in *S.pombe*, many repetitive elements and retrotransposons, priRNA can be generated and target their complementary sense transcripts, thus probing for RNAi dependent heterochromatin establishment (Halic and Moazed, 2010). priRNA are thought to be an RNAi dependent surveillance mechanisms against transposons and other mobile elements, which are known to be transcribed bi-directionally. Antisense transcription would drive the generation of

priRNA and target RNAi to form heterochromatin at these loci. Deletion of the exosome subunit Rrp6 causes an increase in antisense transcription and spurious heterochromatin establishment in those loci in which antisense transcripts are usually degraded (Marasovic et al., 2013). The deletion of the enzyme Triman, responsible for priRNAs biogenesis, counterbalances the lack of Rrp6 and prevents the non-regulated establishment of H3K9 methylation. Moreover, Triman is required for establishment of heterochromatin at centromeres in cells in which Dcr1 and Ctr4 have been re-introduced after deletion. This indicates a fundamental role for priRNAs in inoculating the initial H3K9 methylation and RNAi targeting of centromeric transcripts (Halic and Moazed, 2010; Marasovic et al., 2013).

Another class of siRNA has been recently classified and called primary RNAs. These small RNAs are both Rdp1 independent and Dicer dependent, and helped defining a model for RNAi heterochromatin formation in which convergent transcription, Dcr1 concentration and low transcription efficiency define the loci in which heterochromatin is going to be more likely established (Yu et al., 2014). This is very well supported by the fact that centromeric transcription lacks strong termination and poly-adenylating signals and that generation of low levels of siRNAs is independent of H3K9 methylation (Halic and Moazed, 2010; Yu et al., 2014). It seems therefore that the initiating signal for heterochromatin establishment is determined by the generation of two distinct classes of small RNAs, primal and primary RNAs, which recruit RNAi machinery to convergently transcribed loci.

Heterochromatin at the Mating type (MAT) locus

The silent Mating type (MAT) locus is located on the right arm of Chromosome 2 and spans a region of roughly 30 kb, distant from both centromeric and telomeric loci. Within this region there are three linked loci: *mat1*, *mat2-P* and *mat3-M* (Figure 1.3). *mat1* is the only locus expressed and determines the actual Mating type of the cell, either P or M. *mat2-P* and *mat3-M* are silenced by heterochromatin and are separated from each other by a 11 kb K-region, containing the 4.3 kb *cenH* repeats, that share 96% homology with the centromeric *dg* and *dh* repeats (Grewal and Klar, 1997). In homothallic *h⁹⁰* strains, a gene recombination event converts the gene cassette in *mat1* into one of the two cassettes, *mat2-P* or *mat3-M*, in a preferential fashion: if *mat1-P*, it will recombine with *mat3-M*, if *mat1-M* it will recombine with *mat2-P* (Grewal and Klar, 1997). A specific protein complex, the Recombination Promoting Complex (RPC), containing the proteins Swi5 and Swi2, is responsible for mediating the conversion between mating cassettes in a directional way (Jia et al., 2004). In

homothallic strains this interconversion event happens every second cell division, and it is driven by an unrepaired single strand break during DNA replication that becomes an intermediate for DNA recombination (Holmes et al., 2005; Kaykov et al., 2004). At the MAT locus, two distinct pathways are redundant to establish heterochromatin at the *mat2-P* and *mat3-M* cassettes.

Loss of heterochromatin at the Mating type locus fosters spurious recombination events between cassettes, impairing the normal Switching (Swi) phenotype in homothallic *h⁹⁰* strains and inducing meiotic recombination in the K-region (Lorentz et al., 1992). One of the genes isolated to be responsible for defective MAT type switching is Swi6, encoding for the HP1 homolog. Swi6 is binding through the H3K9me mark and recruiting the SHREC deacetylase complex to establish and spread heterochromatin at the MAT locus (Hall et al., 2002). Establishment of heterochromatin at the MAT locus is dependent on the *cenH* element, which is transcribed and targeted by RNA interference that in turn recruits the CLRC complex to deposit the H3K9me mark (Hall et al., 2002). Replacement of the *cenH* element with an *ura4+* reporter gene resulted in cells having a variegated mating phenotype and epigenetic silencing (Grewal and Klar, 1996). These results indicated that any epigenetic mechanism established for gene control at the MAT locus is heritably transmitted through various cell generations, independently of the DNA sequence. On the other hand, *cenH* genomic integration in other genomic loci is able to establish heterochromatin in a RNAi dependent fashion (Hall et al., 2002). However, RNA interference is dispensable for heterochromatin maintenance and inheritance which is guaranteed by another pathway, dependent on the ATF/CREB transcription factors Atf1/Pcr1 (Hall et al., 2002; Jia, 2004). These DNA binding proteins can recognize two heptameric sequence (called *REIII* element), between the *cenH* repeats and *mat3-M* cassette, and then form complexes with the histone deacetylase Clr3, Clr6 and the HP1 protein Swi6 to limit RNA polymerase II transcription and silence the entire locus (Kim et al., 2004; Sugiyama et al., 2007; Yamada et al., 2005).

Heterochromatin at telomeres

Heterochromatin at telomeres and telomeric associated sequences (TAS, also known as sub-telomeres) is essential to protect chromosome ends from inter- and intra-chromosomal recombination, to maintain genomic stability (Bisht et al., 2008; Schoeftner and Blasco, 2009). Like in the case of MAT locus, RNA interference is involved in heterochromatin formation at sub-telomeric regions (that harbor *cenH* like repeats, Figure 1.3) but it is not the main pathway

directing it. In fact, RNAi mutants do not have heterochromatin defects at telomeres, as the HP1 protein Swi6 maintain its localization and there are no drops in H3K9me levels (Kanoh et al., 2005). Mutations in Swi6, Clr4, the DNA binding protein Taz1 and Rap1 all affect telomeric silencing (Cooper et al., 1997; Kanoh and Ishikawa, 2001; Nimmo et al., 1998; Schoeftner and Blasco, 2009). In the absence of Swi6 in particular, H3K9me is restricted at the *cenH*-like repeats, overlapping with the *cenH* siRNA distribution, indicating a role for RNAi in the establishment of sub-telomeric heterochromatin starting from centromeric *dg* and *dh* like homologous loci. Deletion of the *cenH* like homology region causes an effect similar to the deletion of any RNAi factor, indicating that targeting of sub-telomeric transcripts is required for heterochromatin establishment at sub-telomeres (Kanoh et al., 2005). RNAi components and H3K9 methylation are in fact distributed all over the 40 kb sub-telomeric region (Cam et al., 2005). Taz1 and Rap1 are part of a conserved protein complex called Shelterin, involved in chromosome end protection, to prevent end-to-end inter- and intra-chromosome fusion and to establish silencing across the telomeric locus (Bisht et al., 2008; de Lange, 2009). The telomeric Taz1, homolog of the mammalian Shelterin TRF2, is responsible for recruiting Swi6 through Clr4 activity, to establish de novo H3K9me and induce heterochromatin silencing (Kanoh et al., 2005). Deletion of Taz1 causes an increase in telomeric repeats length (Shelterin is involved in inhibiting telomerase activity once assembled on the chromosome end (de Lange, 2009)) and loss of silencing for telomeric transcripts (Kanoh et al., 2005). Deletion of Taz1 and RNAi factors releases Swi6 from the sub-telomeres, indicating two redundant pathways involved in establishing and spreading heterochromatin at telomeres (Kanoh et al., 2005; Schoeftner and Blasco, 2009). Moreover, Taz1, together with the Shelterin Ccq1 and RNA interference, are fundamental to recruit the SHREC deacetylase complex to telomeres, which deacetylates and adjust nucleosome occupancy to induce silencing (thanks to its subunits Clr3 and the chromatin remodeler Mit1)(Sugiyama et al., 2007). Telomeric repeats in *S.pombe* are transcribed, even though the role of transcription for telomere length regulation, chromosome end protection or silencing is still obscure (Bah et al., 2012). Recent studies reveal novel roles for telomeres in regulating centromeric heterochromatin assembly. The rate limiting distribution of heterochromatin factors through the different heterochromatin domains makes cells lacking normal Shelterin assembly able to bypass RNAi for pericentromeric heterochromatin formation. This would be caused by an altered recruitment of silencing factors to telomeres and their re-distribution to centromeres (Tadeo et al., 2013), indicating a cross-talk between different chromatin domains and explaining also the reason for clustering in similar nuclear compartments. Further confirmation of this co-regulation is found in the need for proper

Introduction

telomere localization (through the Bouquet complex anchoring to the nuclear membrane) to re-assemble pericentromeric heterochromatin in meiosis, which is required for CENP-A deposition and Spindle Pole Body attachment (Klutstein et al., 2015).

Heterochromatin formation main players: HP1 and chromodomain proteins

Silencing is mediated by the coordinated action of different proteins and chromodomain containing proteins play a fundamental role in many Eukaryotic organisms to ensure the establishment and propagation of heterochromatin. The first evidences of chromodomain proteins having a fundamental role for heterochromatin stability were from studies done in *Drosophila melanogaster*, as previously mentioned.

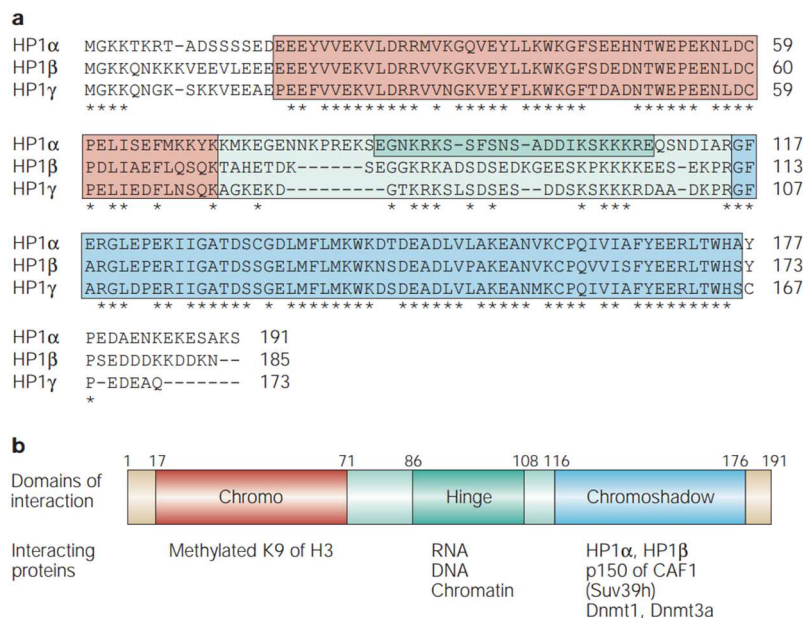


Figure 1.5: HP1 proteins and their domains

From (Maison and Almouzni, 2004): (a) alignment between the three mammalian HP1 isoforms: HP1 α , HP1 β and HP1 γ . While the three Chromodomains (CD, in red) and Chromoshadow domains (CSD, in blue) share a high degree of conservation, the Hinge region is characterized by being not conserved and positively charged with many basic residues. (b) Domain organization of the mammalian HP1 α protein with its interaction partners. In red, the CD, in green the Hinge domain and in blue the CSD. This organization is shared among HP1 proteins.

In 1986, James and Elgin discovered and characterized through a screening for PEV (Position Effect Variegation) suppressors, the HP1a (Heterochromatin Protein 1a, encoded by the *Su(var)2-5* gene, reviewed in (Eissenberg and Elgin, 2014)), a non-histone protein associated with heterochromatin (James and Elgin, 1986). Later, more evidences showed that the HP1a gene was associated with the control of PEV: a G to A transition mutation causing a splicing defect of the gene was responsible for a strong dominant PEV suppressor phenotype (Eissenberg et al., 1990) and PEV increase could be observed when over-expressing the HP1a

Introduction

(under the Hsp70 heat inducible promoter) after transduction of a P-element, carrying the HP1a gene, transformed into fruit-fly germline cells (Eissenberg et al., 1992). This initial studies were highlighting the role of HP1 chromodomain proteins in establishing and maintaining heterochromatin domains and tightly defining their borders. Many HP1 homologs are found in different organisms, from fission yeast to humans, with isoforms having distinct functions and interacting with different complexes (Lomberk et al., 2006a).

In the budding yeast *S. cerevisiae* silencing does not occur through deposition of the H3K9 methyl mark and its own read out (there is no H3K9 methylation or HP1 homologs), but through coupling of SIR mediated histone H4K16 de-acetylation and SIR complex binding. The SIR complex, formed by the deacetylase Sir2, the Sir3 and Sir4 proteins, recognize specifically de-acetylated nucleosomes, binds to them through the Sir3 BAH domain and oligomerize through its Sir3 and Sir4 subunits, thereby spreading the silencing from the recruiting locus (recruitment occurs through protein-protein interaction with DNA binding factors such as ORC complex, Abf1 and Rap1. The sequence elements required for the recruitment are called E –essential- or I –important- silencers). The H3K79 histone globular domain modification antagonizes the SIR complex binding to chromatin, thus marking actively transcribed genes (Oppikofer et al., 2013).

In a similar way as the SIR complex in *S. cerevisiae*, HP1 recognition of the H3K9 histone post-translational modification is essential for establishing a silencing platform, necessary for the recruitment of different complexes to repress transcription. This recognition occurs through the chromodomain, first discovered by studying the *Drosophila* Polycomb gene, which is showing an N- terminal region having high homology with the HP1 N- terminal region (Paro and Hogness, 1991) (Figure 1.5). The Polycomb-group proteins chromodomain (there are at least 5 different variants of CBX proteins in the canonical mammalian PRC1 complex, with different affinities for the H3K27, some of them showing dual affinity for H3K9 methylation (Bernstein et al., 2006; Di Croce and Helin, 2013)) recognizes specifically H3K27 methylation, a mark of facultative heterochromatin, while sharing structural and sequence similarity with the HP1 chromodomain (Min et al., 2003). However, swapping of the Polycomb chromodomain with HP1 chromodomain alters the localization and function of these factors to chromatin, indicating specialized roles for these domains in dictating the distribution of heterochromatin proteins (Fischle et al., 2003). The chromo- (Chromatin Organization Modifier) domain is found in all HP1 proteins and in many chromatin specific factors, with a broad range of affinity for different target histone methylation marks as reviewed in (Eissenberg, 2012). Mutations in the chromodomain, that alter its specificity or affinity for its

target H3K9 methylation mark, impair the silencing activity of HP1 proteins, revealing the importance of H3K9 recognition to mediate heterochromatin establishment (Eissenberg, 2012; Platero et al., 1995). Moreover, post-translational modifications such as Ser-83 phosphorylation of the fruit-fly HP1 gamma, can alter the localization of HP1 proteins from heterochromatin to euchromatin, indicating the existence of a HP1 sub-code beside the recognition of the H3K9 heterochromatin mark (Lomber et al., 2006b). Protein of the HP1 family share other common features beside the chromodomain, such as the presence of a widely unstructured part called Hinge domain and the Chromo-Shadow (CSD) domain (Figure 1.5). The Chromo-Shadow domain was first identified through its similarity with the Chromodomain (Assland and Stewart, 1995). Structure and sequence analysis revealed the presence of a conserved pentameric motif, the PXVXL motif, necessary for CSD dimerization (Smothers and Henikoff, 2000). The ability of the CSD to dimerize is required to recruit more HP1 proteins to chromatin and establish heterochromatin silencing, coupled with H3K9me recognition from the chromodomain (Thiru et al., 2004).

The Hinge domain separates the Chromodomain from the Chromo-Shadow domain and represents the most flexible part of the protein, with the least conserved sequence (Munari et al., 2013; Nishibuchi and Nakayama, 2014). Several studies (Keller et al., 2012; Mishima et al., 2013; Muchardt et al., 2002) showed that the Hinge region binding to RNA and DNA was essential to coordinate HP1 localization to chromatin, along with H3K9 methyl mark recognition. Particularly, the Hinge of the *S. pombe* HP1 Swi6 was shown to specifically bind to centromeric RNAs and a model of coordinated eviction of Swi6 from chromatin was proposed: Swi6 binds centromeric RNAs to prompt them for degradation. To do so, Swi6 would dissociate from heterochromatin and interact with transcribed RNAs with its Hinge domain (Keller et al., 2012).

Other proteins of fundamental importance for heterochromatin formation are the Su(var)3-9 methyltransferases (such as the *S.pombe* Clr4). First characterized in *Drosophila* (Tschiersch et al., 1994), this class of enzymes has also a chromodomain. The chromodomain allows them to couple directly the deposition of H3K9 methylation with its read-out, facilitating the spreading of the histone mark, independently of the underlying DNA sequence (Audergon et al., 2015; Rangunathan et al., 2014).

The chromodomain: structure and function in heterochromatin regulation

The Chromatin Organization Modifier domain or Chromodomain (CD) is a small domain of

roughly 10 kDa, formed by a three antiparallel β -strand and a C-terminal α -helix. It is one of the first nucleosome binding domains characterized in details, as first studies performed showed it had a high specificity for the histone tail H3 methylated at the Lysine 9 residue (Bannister et al., 2001; Lachner et al., 2001). Confirmation for this binding specificity came from additional *in vivo* and *in vitro* experiments, showing a strict dependence between chromodomain interaction with nucleosomes and the presence of the H3K9 methylation mark. In fact loss of the H3K9me causes HP1 proteins delocalization from chromatin, in yeast, *Drosophila* and mammalian cells (Eissenberg, 2012; Jacobs et al., 2001; Lachner et al., 2001; Nakayama et al., 2001; Schotta et al., 2004). The first structure with the specifically recognized target (Jacobs and Khorasanizadeh, 2002), the H3K9 methylated histone tail, highlighted that the residues interacting with the modified Lysine 9 form an aromatic cage around it, with affinity extremely sensitive to the degree of methylation (either mono- di or tri- methyl state). The three residues of the aromatic pocket (in Chp1CD: Tyrosine Y22, Tryptophan W44 and Tyrosine Y47) interact with the positively charge methyl-lysine residue through the interaction of the cationic charge of the ϵ -amine group with the delocalized electron cloud of the Tryptophan aromatic ring (cation- π interaction)(Hughes et al., 2007). The degree of methylation changes the cationic strength of the ϵ -amine, with the tri-methyl state being more positively charged and having the strongest binding affinity *in vitro* (Figure 1.6). While chromodomains are structurally conserved among the HP1 and Polycomb family of proteins, they are able to distinguish with high specificity between H3K9 methylation and H3K27 methylation (Fischle et al., 2003; Min et al., 2003). As already mentioned, swapping of the HP1 CD with the Polycomb CD causes a mis-localization of the HP1 protein, with consequent loss of heterochromatin (Platero et al., 1995). This indicates that CDs not only are able to read the Lysine residue that is methylated but also the surrounding residues along the histone tail, conferring further specificity to the binding. The deposition of different methyl- marks therefore influences the distribution of chromodomain proteins with different specificities, as in case of the HP1 with the H3K9me and Polycomb with H3K27me histone marks. However, some CDs, such as the mammalian Polycomb isoforms Cbx2 and Cbx7, share similar affinities for H3K9 and H3K27 methylation (Bernstein et al., 2006). This hints on the possibility that different Polycomb isoforms contribute to the recognition of the H3K27 or H3K9 methylation mark in different ways, thus modifying the efficiency of recruitment or the nuclear distribution of the Polycomb complex in a regulated way.

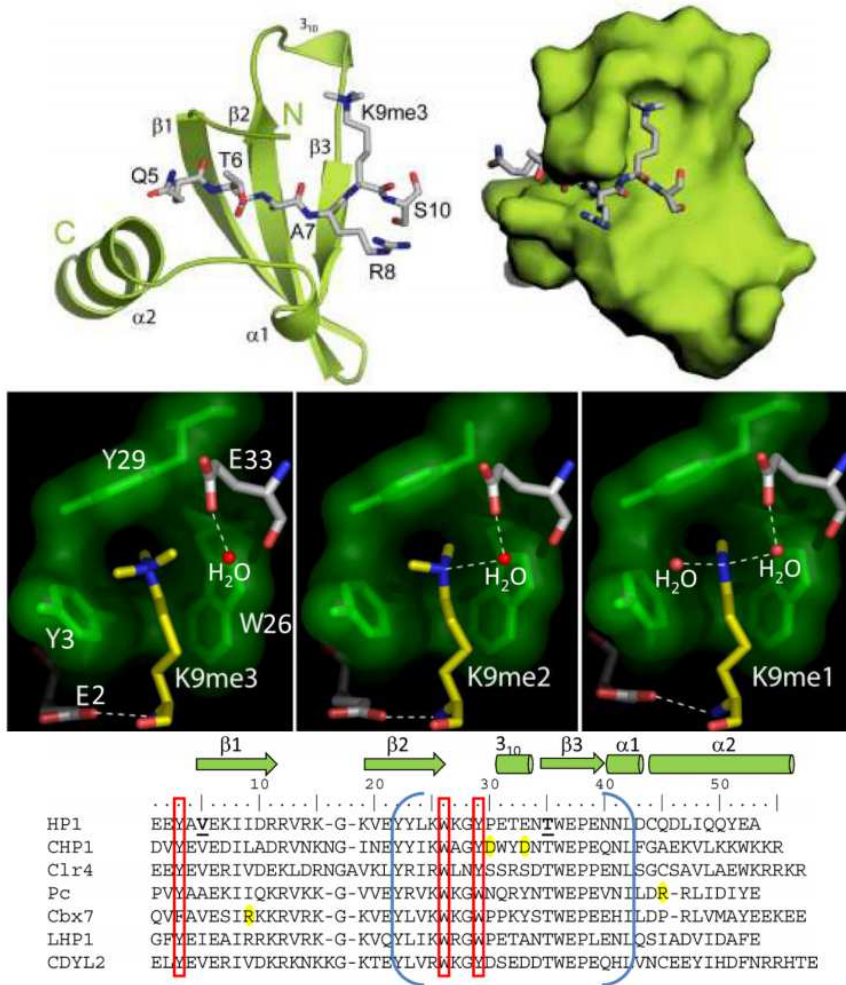


Figure 1.6: The chromodomain (CD) specifically recognizes H3K9 methylation

From (Blus et al., 2011): X-ray crystal structure of the canonical chromodomain (in green) of the HP1 protein, bound to the H3K9me3 peptide (in grey). The cation- π interaction is mediated by three aromatic residues (two conserved Tyrosine and a Tryptophan, Y22, W44 and Y47 in the *S.pombe* Chp1) with their aromatic rings delocalized electrons stabilizing the positive charge of the methylated lysine residue (H3K9me) of the histone tail. Different methylation states alter the charge on the lysine, thereby changing the affinity of binding. As shown in the alignment, the three residues (in the red boxes) are conserved throughout evolution, from fission yeast (Chp1 and Ctr4) and *Drosophila* (Pc and HP1) to mammals (Cbx7, CDYL2).

CD recognition of its target can be regulated in two ways: 1) by modifying the substrate recognition sequence or 2) by post-translationally modifying the chromodomain itself.

Growing evidences show that the histone tail can be modified to prevent binding of HP1 proteins *in vivo*. Histone H3S10 is phosphorylated by the Aurora B kinase/Ikp1 upon entry in S phase (Hsu et al., 2000). A study by Hirota and colleagues in 2005 showed that such a modification occurs on histone tails harboring already the H3K9 methyl mark, and that the dual marking exists *in vivo*. The binding affinity of HP1 proteins for the H3K9me in the presence of Serine 10 phosphorylation is severely impaired, and deposition of the H3S10phospho mark causes the delocalization of HP1 proteins from chromatin (Hirota et al., 2005). Both histone H3K9 methylation and H3S10 phosphorylation can be reversible (PP1 phosphatase reverts

H3S10 phosphorylation (Hsu et al., 2000)), showing that HP1 recruitment to chromatin is subjected to a complex layer of regulation involving substrate modification. Different H3 Lysine residue that have been shown to be methylated *in vivo* are adjacent to Threonine or Serine residues that can be phosphorylated. It is the case for Lysine 4 (H3K4), close to Threonine 3 (H3T3) and Lysine 79 (H3K79) close to Threonine 80 (H3T80). Both mark of actively transcribed regions, H3T3 has been shown to be phosphorylated *in vivo* by the Haspin kinase and to lower the binding affinity of the CHD1 CD, regulating its binding in a similar manner as HP1/H3S10phospho (Dai et al., 2005; Flanagan et al., 2005).

Moreover, phosphorylation by the DYRK1A kinase at residues H3T45 and H3S57 affects the binding of the three mammalian HP1 isoforms, HP1 α , HP1 β and HP1 γ (Jang et al., 2014).

It is still unclear the effect of direct post-translational modifications on the CD on its binding activity *in vivo*. It is well known that HP1 proteins are heavily phosphorylated by Casein kinase (CKII) in *Drosophila* and fission yeast (Shimada et al., 2009; Zhao and Eisenberg, 1999; Zhao et al., 2001) and that this is required for proper silencing and heterochromatin formation. There is still little evidence of regulated CD modifications that alter the binding affinity *in vivo*: a study from Ayoub et al. showed that HP1 β CD is phosphorylated on residue T51 in response to DNA damage, lowering the CD binding affinity for H3K9me and therefore dissociating HP1 β from chromatin (Ayoub et al., 2008). Interestingly, phosphorylation of Serine 42 of the chromodomain of the mouse Polycomb Cbx2 protein lowers its affinity for the H3K9 methylated peptides while increasing the binding for H3K27 methylated peptides (Hatano et al., 2010). Post-translational modifications could therefore regulate not only the binding affinity of the chromodomains but also their specificity.

Another striking feature of CDs is their ability to interact with the nucleic acids. The first example of such interaction was discovered in *Drosophila*, with the chromodomain of the MOF protein, member of the MSL (Male Sex Lethal) complex, responsible for X inactivation and dosage compensation (Akhtar et al., 2000). MSL is a riboprotein complex with two structural RNAs, roX1 and roX2 and two chromodomain proteins MOF and MSL3. Both chromodomains were found to interact with the roX2 RNA, and MOF dependent acetylation of MSL3 is required to stabilize the binding (Buscaino et al., 2003). The interaction with the roX2 RNA is essential for the stability and function of the MSL complex.

MSL3 CD recognizes specifically H4K20 mono-methylation but DNA binding is essential to mediate the association, forming a stable ternary complex (DNA, MSL3, H4K20 tail) (Kim et al., 2010). Association of chromodomains with both DNA and RNA has been shown now in

several organisms to be likely relevant *in vivo*. The chromodomain of the mouse Polycomb Cbx7 requires RNA binding to mediate its recruitment to the inactive X chromosome (Bernstein et al., 2006). The human CBX7 is able to bind both the H3K27me histone tail and the ANRIL non-coding RNA and this contributes to control the expression of the tumor suppressor locus INK4b/ARF/INK4a (Yap et al., 2010).

Similarly, the CD of Chp1, member of the RITS complex of *S. pombe*, has been shown to have nucleic acids binding *in vivo*, with the binding interface restricted to the C-terminal positively charged α -helix (Ishida et al., 2012). Swapping of the C-terminal residues of Chp1 CD with the homologous residues from Swi6 CD (which does not bind RNA or DNA *in vitro*) causes transcripts up-regulation and diminished levels of H3K9me at centromeres, hinting on the importance of nucleic acids binding for heterochromatin formation (Ishida et al., 2012). Similarly, Clr4CD binds the RNA *in vitro* (Ishida et al., 2012) even though the importance of this interaction related to H3K9me deposition is still to be explored.

The chromodomain proteins of *Schizosaccharomyces pombe* in heterochromatin formation

There are four chromodomain proteins involved in heterochromatin formation in *S. pombe*: two of them belonging to the HP1 family, namely Chp2 and Swi6, the histone methyltransferase Clr4, homologous to the Su(var)3-9 and the member of RITS complex, Chp1 (Figure 1.7).

Swi6 is the HP1 homolog, involved in heterochromatin assembly at centromeres, telomeres and at the MAT locus (Ekwall et al., 1995). Chp2 is its functional isoform, as well localizing at all three heterochromatin loci. Both Swi6 and Chp2 are required for the assembly of repressive heterochromatin, but Chp2 is less abundant and has a more defined role in recruiting the SHREC (Chp2 has higher affinity than Swi6 for SHREC deacetylase Clr3) and Epe1 to chromatin (Motamedi et al., 2008; Sadaie et al., 2008). Cellular fractionation experiments showed that the interaction of Chp2 with chromatin is more dynamic than that of Swi6 (Sadaie et al., 2008), with more protein concentrated in the nuclear soluble fraction rather than the chromatin bound.

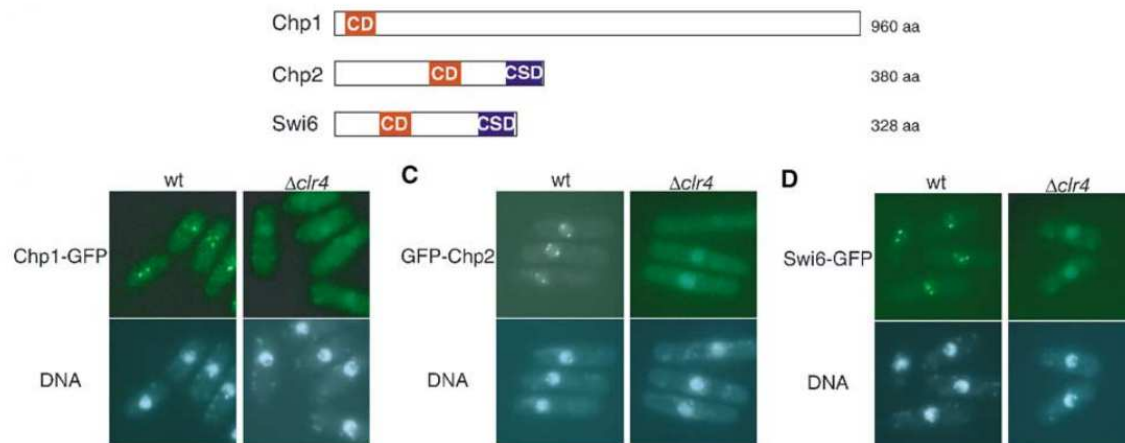


Figure 1.7: Chromodomain proteins in fission yeast

From (Sadaie et al., 2004): Schematic representation of three of the four heterochromatin protein in *S.pombe*: Chp1, Swi6 and Chp2. The fourth one is the methyltransferase Clr4, responsible for the H3K9me mark deposition (not shown in the cartoon, it harbors a CD and a methyltransferase SET domain). Chp1 is member of the RITS-Argonaute complex, whereas Swi6 and Chp2 are fission yeast HP1 homologs. The Chromodomain is highlighted in red (CD), the Chromoshadow domain (CSD) in blue. In the panel below, GFP tagged version of Chp1, Swi6 and Chp2 localize as nuclear foci in *clr4*⁺ cells. In the absence of H3K9me (dependent on Clr4) the proteins are not anymore localized in distinct foci. DNA was stained with Hoechst33342 as cellular control.

Moreover, Swi6 interacts with centromeric RNAs *in vivo* and this has been proposed to be a mechanisms to facilitate RNAi targeting of centromeric transcripts (Motamedi et al., 2008) or their degradation (Keller et al., 2012). Chp2, on the other hand, recruits the SHREC to deacetylate H3K14 (through its Clr3 subunit) and, in cooperation with Swi6, limits the access of RNA Polymerase II to centromeric loci (Fischer et al., 2009; Motamedi et al., 2008). Single Swi6 or Chp2 deletion mutants have similar effect on centromeric transcripts up-regulation, while *swi6Δchp2Δ* double mutant shows a cumulative defect and an increased loss of silencing. This indicates that Swi6 and Chp2 have different non-overlapping roles in heterochromatin formation and that they are involved in distinct silencing pathways (Motamedi et al., 2008). In both cases, however, CSD dimerization mutants (Swi6 L315E and Chp2 I370E) that are unable to associate through CSD interaction, show a strong impairment in heterochromatin maintenance, highlighting that the two HP1 need self-association to promote silencing (Sadaie et al., 2008). Despite sharing similar overall organization, Swi6 and Chp2 domains are not equivalent in terms of function. Swapping of Swi6 and Chp2 CD, Hinge or CSD domains causes a loss of heterochromatin silencing (Sadaie et al., 2008). When swapping the N-terminal part of the proteins, however, Chp2 did not have a visible loss of silencing, whereas Swi6 had a huge *Kint2::ura4*⁺ reporter transcript upregulation. This is consistent with the role of the N-terminal domain of Swi6 (and similarly the one of the human homolog HP1α), which is heavily

phosphorylated by CKII *in vivo*, in regulating its function (Nishibuchi et al., 2014). Localization of both HP1 proteins to heterochromatin loci is dependent on Clr4, as shown by Swi6 and Chp2 ChIP sequencing data and immunofluorescence experiments (Sadaie et al., 2004, 2008). As already mentioned, the methyltransferase Clr4 is depositing the H3K9 methylation mark, which is necessary for the recruitment of HP1 proteins to heterochromatin (Nakayama et al., 2001) (Figure 1.7). The catalysis of H3K9me is mediated by the SET domain of Clr4, the sole H3K9 methyltransferase identified in fission yeast (Rea et al., 2000). The name SET domain derived from the Su(var)3-9 and Enhancer of Zeste proteins, where it was first characterized. Clr4 has its own CD, with which it can recognize H3K9 methylation and methylate nearby nucleosomes, in a read-write mechanism (Al-Sady et al., 2013). This is particularly important to epigenetically maintain the epigenetic mark, even in the absence of RNAi or other heterochromatin initiating stimuli (Audergon et al., 2015; Ragunathan et al., 2014). The CD of Clr4 binds with higher affinity H3K9 trimethylation (H3K9me3) rather than the H3K9 dimethylation (H3K9me2), and its SET domain catalyzes with more efficiency H3K9me2 deposition (which is the most abundant form of H3K9 methylation in *S.pombe*). This has a huge functional importance, as reported in (Al-Sady et al., 2013), since the HP1 protein Swi6 CD has higher affinity than Clr4CD for H3K9me2 but not for H3K9me3. In the competition between the two proteins at heterochromatin, Swi6 will outcompete Clr4 for H3K9me2 binding, given the abundance of this methylation state. On the other hand, Clr4 may use the fewer H3K9me3 spots as anchor site for H3K9me deposition. Swapping of Clr4CD with a Chp1CD mutant showing higher affinity for H3K9me2, while maintaining the same affinity for H3K9me3 (Chp1CDF61A as shown in (Schalch et al., 2009)), causes a huge defect in spreading the H3K9 methylation mark beyond the initiation point at the MAT locus (Al-Sady et al., 2013). This indicates that the selective recognition from Clr4CD between different H3K9 methylation states is necessary to coordinate its function with HP1 proteins. To sum up, this is an example of how the localization of different chromatin readers, recognizing the same histone mark, is organized by the individual CD affinities for H3K9me and different methylation states.

The fourth CD protein involved in heterochromatin formation and focus of the present thesis work, is Chp1. Chp1 is a member of the RNA induced transcriptional silencing complex (RITS), necessary to anchor Argonaute to centromeric heterochromatin, as previously discussed (Verdel et al., 2004) . It is a relatively big protein, with a size of 108.73 kDa, interacting with the N-terminal part of the GW protein Tas3 with its C-terminal domain (Schalch et al., 2011). Together, Chp1 and Tas3 form a functional sub-complex of the RITS

Introduction

that binds H3K9me marked nucleosomes and form a platform for recruiting Argonaute (loaded with mature siRNAs) to centromeric transcripts. A Tas3 mutant no longer able to bind to Chp1 had a severe loss of centromeric heterochromatin ((Tas3 Δ (10-24))), indicating the importance of Chp1-Tas3 integrity for RITS formation and RDRC complex recruitment (Debeauchamp et al., 2008). Chp1 and Tas3 were shown to interact also independently of Argonaute in *ago1* Δ cells, though not recruited to heterochromatin loci (that are depleted in H3K9me in the absence of RNAi) (Petrie et al., 2005).

Chp1 is essential for heterochromatin establishment and maintenance at centromeres, whereas at telomeres and at the mating type loci *chp1* Δ cells do not show evident defects in silencing (Sadaie et al., 2004). This is explained by the presence of additional pathways, which act in redundancy with RNAi to maintain heterochromatin at telomeres and the MAT locus, as discussed previously.

However, establishment of heterochromatin is dependent on Chp1 at all three heterochromatin loci (Sadaie et al., 2004). As shown by Cam and colleagues, RNA interference components recruitment to all three heterochromatin loci is dependent on Clr4 deposition of H3K9 methylation (Cam et al., 2005). The H3K9 methylation serves as anchoring point not only for HP1 proteins, but also for Chp1-RITS, and high affinity of Chp1 CD for H3K9me is necessary to establish heterochromatin (Schalch et al., 2009).

The specific interaction of Chp1 with its target histone modification is therefore a key feature to establish and maintain heterochromatin in *S. pombe* and the focus of the work presented in this thesis.

Aim of the thesis

Heterochromatin formation is fundamental for the regulation of Eukaryotic genome expression and spatial organization inside the Nucleus. Heterochromatin is characterized by the presence of a specific histone post-translational modification, the H3K9 methylation, which signals for the establishment and maintenance of gene silencing. The binding of this histone mark is operated by specific chromatin readers called chromodomain proteins, which mediate the initial interaction with H3K9me and then recruit additional complexes responsible for H3K9me deposition and chromatin compaction (Eissenberg, 2012; Eissenberg and Elgin, 2014; Nakayama et al., 2001; Nishibuchi and Nakayama, 2014). Many structures of chromodomains were published to date, but mostly either CDs alone or binding to H3K9me peptides, mimicking the methylated histone H3 tail (Fischle et al., 2003; Horita et al., 2001; Ishida et al., 2012; Jacobs and Khorasanizadeh, 2002; Nielsen et al., 2002; Schalch et al., 2009). No structural information is available on the binding of chromodomains with the H3K9me3 Nucleosome, their natural interactor. How chromodomain proteins interact with the H3K9me Nucleosome? Chromodomain containing proteins have been shown to have different functions: among them are HP1 proteins (Nishibuchi and Nakayama, 2014), structural components of heterochromatin, H3K9 methyltransferases, such as the Su(var)3-9 family (Tschiersch et al., 1994) and chromatin remodelers. How all these functions are regulated at the same genomic loci, despite the recruiting signal being the same H3K9 methylation? Moreover, different chromodomains have been shown to be able to bind to nucleic acids (Akhtar et al., 2000; Bernstein et al., 2006; Buscaino et al., 2003; Ishida et al., 2012). How does this interaction regulate their physiological function?

The aim of the work presented in this thesis was to understand the mechanism of binding of chromodomain proteins to chromatin and to shed light on how this interaction might be regulated in the context of heterochromatin formation. Extensive studies have been performed on heterochromatin and silencing in the model organism *S. pombe*. The attention of this Thesis work was focused on nuclear *S. pombe* chromodomain proteins, and in particular on Chp1, member of the RITS-Argonaute complex (Verdel et al., 2004), to investigate how its interaction with the H3K9me Nucleosome influences heterochromatin establishment and maintenance. Biochemical and *in vivo* experiments were employed to give an answer to all these important questions in the chromatin field.

RESULTS

In vitro reconstitution of the Chp1CD-H3K9me3 Nucleosome complex

In order to gain insights on how chromodomains interact with nucleosomal particles, the Chp1CD-H3K9me3 Nucleosome complex was reconstituted *in vitro*, starting from the single components. The chromodomain of Chp1 was expressed and purified from the bacterial *E.coli* BL21(DE3)pLysS strain from the pET28a expression vector as an His₆-Thrombin digestion site-SUMO tag fusion construct (Schalch et al., 2009). The expressed SUMO-Chp1CD was purified by Ni-NTA chromatography and run at around 25 kDa on a 15% SDS-PAGE gel (Figure 2.1 A). Single *Xenopus laevis* histones (H2A,H2B,H3,H4) were purified as well from the bacterial *E.coli* BL21(DE3)pLysS expression strain from the pET3 expression vector as previously described (Dyer et al., 2004). Histones were DMSO extracted from isolated inclusion bodies and purified by gel filtration under denaturing conditions (Dyer et al., 2004). After gel filtration, histones were dialyzed in 3 mM β-mercaptoethanol before being lyophilized for long term storage. Lyophilized histone H3 was then re-suspended in Alkylation buffer and chemically modified by addition of (2-bromoethyl) trimethylammonium bromide to generate the Methyl-lysine analog (MLA) mimicking the H3K9 tri-methylation. The procedure installing MLA on histone H3 was done strictly following a previously published protocol (Simon et al., 2007). The modifications obtained are exactly equivalent to H3K9 methylation and have been reported to be recognized by HP1 and Polycomb proteins with the same affinity as their native counterpart (Simon et al., 2007). To check for proper installation of the MLA mark we performed mass-spectrometry (data analysis by Dr. Thomas Fröhlich – Gene Center – Munich) and western blot on the modified H3 histones. The mass spectrometry profile gives one clear peak at the expected mass of 15297 Da (Figure 2.1 D). The additional peak seen at around 15340 Da has been described as an artifact, present in the original batch of histone H3 before alkylation, not reflecting a true contaminating specie (Simon et al., 2007). The western blot (Figure 2.1 E) confirmed that not only the MLA modification was installed correctly on histone H3, but also that an antibody elicited against the natural H3K9me3 mark could recognize it as well. This indicates the biochemical and functional equivalence between the MLA system and the *in vivo* H3K9 methylation. From this point onwards the MLA H3K9me3 Nucleosomes will be referred as H3K9me3 Nucleosomes.

Results

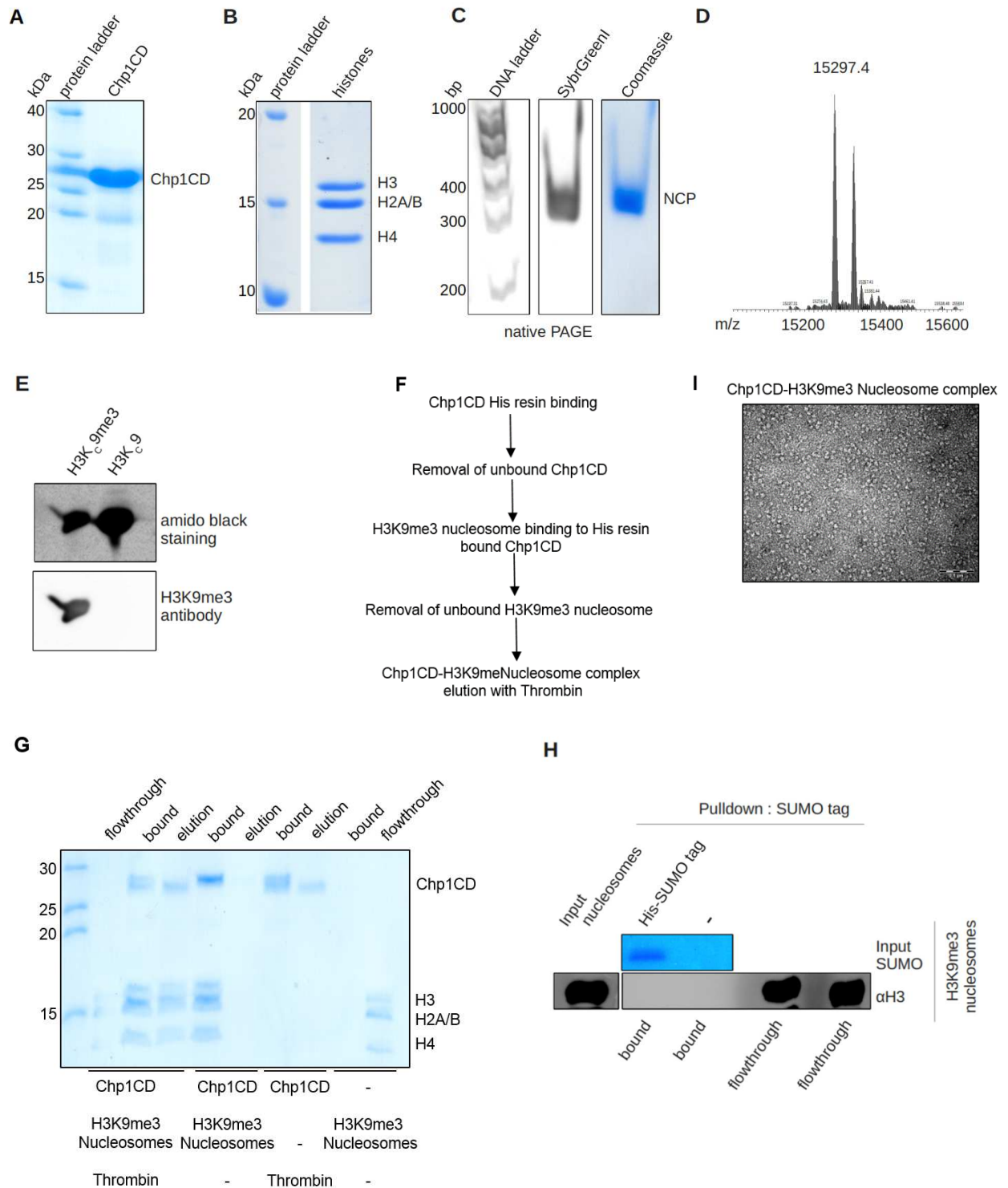


Figure 2.1: In vitro assembly of the Chp1CD-H3K9me3 Nucleosome complex

*A) Coomassie staining of a 15% SDS-PAGE gel showing the Chp1 Chromodomain (Chp1CD) purification. Chp1CD runs above 25 kDa (7 kDa Chp1CD + SUMO tag). (B) In vitro reconstitution of the nucleosome octamer run on a 15% SDS-PAGE gel: all *Xenopus laevis* histone bands are visible. H2A and H2B cannot be separated and run as duplet. (C) Nucleosome reconstitution: nucleosomes were run on a 5% TBE-Native acrylamide gel and stained with SyBr GOLD and Coomassie. (D) Mass spec analysis performed on H3K9me3 histones (data provided by Thomas Fröhlich – Gene Center). A main peak at the mass of 15297.4 Da indicates that the alkylation reaction was successful. The +45 Da extra peak was already described as uninfluential for the purity of the*

Results

preparation (Simon et al., 2007). (E) Western blot on the H3 MLA histones. The anti-H3K9me3 antibody (ab8898 Abcam) is specific to the alkylated histones and does not react with the untreated histone H3. F) Experimental scheme for the Chp1CD-H3K9me Nucleosome complex assembly. G) Coomassie stained 15% SDS-PAGE gel showing the assembly of Chp1CD-H3K9me3 Nucleosome complex and its elution from the Ni-NTA resin by Thrombin digestion. H) Nucleosome pull down/western blot experiments using the His₆-SUMO tag alone bound to the Ni-NTA resin. As expected, the SUMO tag cannot interact with H3K9me3 Nucleosomes. I) Negative stain EM micrograph showing the Chp1CD-H3K9me3 Nucleosome complex.

The reconstitution of the Nucleosome core particles (NCP) was performed starting from lyophilized histones and strictly according to the Luger lab procedure (Dyer et al., 2004; Luger et al., 1999). Histones were re-suspended in denaturing Guanidium-HCl unfolding buffer (see Materials and Methods Table 4) and mixed with the following molar ratio (H2A:H2B:H3:H4 1.1 : 1.1 : 1 : 1) for Nucleosome octamer reconstitution. The assembled histone octamer was then purified by gel filtration and its purity and integrity was checked on SDS-PAGE (Figure 2.1 B). The NCP was completely reconstituted by mixing the 601 Widom 147 bp DNA (Lowary and Widom, 1998) and the assembled octamer (DNA:Octamer 1.1 : 1 ratio) and by performing a step gradient dialysis with decreasing salt concentration. Nucleosome reconstitution was assessed by Native-TBE PAGE (Figure 2.1 C): the NCPs can be stained both with SyBr GOLD, specific for nucleic acids, and with Coomassie Blue. Moreover, EMSA assays on the same Native PAGE-TBE gels showed that NCP particles run slower than the free 601 147bp DNA sequence, indicating that the octamer was stably assembled onto the DNA. To assemble the NCP, the 601 Widom sequence was preferred over other possible DNA sequences (such as the 5S RNA gene of *Lytechinus variegatus* (Dyer et al., 2004) due to its reported high stability (Lowary and Widom, 1998): this prevented the need of subsequent steps of heat shifting of the NCP particle to center the octamer onto the 147 bp DNA sequence and the risk of having a non-homogenous population of Nucleosome particles with different octamer-DNA orientations. Moreover, NCP particles assembled with the 601 Widom sequence have been previously crystallized, showing that they can be obtained with high purity and are stable for structural studies (Vasudevan et al., 2010). The Chp1CD-H3K9me3 Nucleosome complex was assembled with a pull down strategy, using the His₆-SUMO tag to bind the Chp1CD on the Ni-NTA resin. After anchoring the Chp1CD onto the Ni-NTA resin, we added the H3K9me3 Nucleosomes and incubated on ice. Elution was performed by adding Thrombin, which digested between the His₆ tag and the SUMO tag. Thrombin addition caused a shift in size of the Chp1CD protein (fourth lane from the left, Figure 2.1 G), due to the release of the His₆ tag, with consequent elution of the SUMO-Chp1CD-H3K9me3 Nucleosome complex. Without addition of Thrombin the complex stayed bound to the resin (fifth lane from the left, Figure 2.1 G). The Ni-NTA resin alone incubated with H3K9me3 Nucleosomes did not show any non-

specific NCP particle binding (ninth lane from the left, Figure 2.1 G): this indicated that the complex was indeed formed due to the specific recognition by the Chp1CD of H3K9me3 Nucleosomes. To further investigate the interaction, the SUMO tag was tested for the binding with H3K9me3 Nucleosomes. The Nucleosome pull down assay was performed using only the SUMO tag, with the same experimental procedure applied for the entire SUMO-Chp1CD construct. Given the higher sensitivity compared to the classical Coomassie Blue staining, Western blot using an anti-H3 specific antibody was used to test the interaction. SUMO was not able to affinity purify H3K9me3 Nucleosomes (second lane from the left, Figure 2.1 H). It could be therefore excluded that the SUMO-Chp1CD-H3K9me3 Nucleosome interaction was influenced by the presence of the SUMO tag in addition to the Chp1 chromodomain. Once assembled and eluted, the SUMO-Chp1CD-H3K9me3 Nucleosome was checked under the Morgagni (FEI Company) transmission electron microscope by Negative staining Electron Microscopy (EM). The eluted complex (fourth lane from the left, Figure 2.1 G), appears to be homogenous in shape and concentrated enough to cover large part of the carbon coated EM grid uniformly (Figure 2.1 I). This prompted us to further continue with the study, aiming to solve the structure of the Chp1CD-H3K9me3 Nucleosome complex by using cryo-EM.

Chp1CD binding to the H3K9me3 Nucleosome depends on the stability of its interacting loops

A preliminary cryo-EM analysis performed in our laboratory (data not shown), showed the Chp1CD-H3K9me3 Nucleosome complex at a sub-nanometer resolution. The fitting of the crystal structure of the Chp1CD in one of the best resolved cryo-EM density maps showed that the Chp1CD assumed a distinct orientation, with the first loop of the three stranded antiparallel β -sheet making a strong contact with the histone H3 loop (residues 77-81), as well as other minor contacts between the second loop and the C-terminal α -helix and histones H2B and H4. Therefore, the preliminary structural analysis showed that Chp1CD interacts with the Nucleosome core, mainly at histone H3, through the first two loops of its β -sheet, while the C-terminal α -helix protrudes outwards from the Nucleosome surface. In order to prove the described structural model, mutations were done by substituting those residues in the first loop (LOOP1, residues 31-38aa) and in the second loop (LOOP2, residues 48-52aa) of Chp1CD which were close to the Chp1CD contact points with the Nucleosome core in the cryo-EM maps (data not shown). LOOP1 mutant harbors three different mutations: R31S, N33A and N35A (Figure 2.3 A). LOOP1 mutants can be divided in two classes, namely LOOP1A (having

Results

all three mutations, R31S, N33A and N35A) and LOOP1B (having only N33A and N35A mutations). LOOP2 mutants on the other hand can be divided into LOOP2A (harboring an N52A mutation) and LOOP2B (harboring W49A, Y50A and D51A mutations). The mutated residues are conserved in a number of other chromodomains (Figure 2.2 A). In particular residues R31, D51 and N52 are conserved in different chromodomain proteins in *S.pombe* as well as other Eukaryotes, whereas LOOP1 N33 and N35 are found only in a subset of other CDs (N35 is conserved in Clr4, not shown in the alignment in Figure 2.2 A). All the residues in LOOP1 that were mutated are actually aligning with the *S. cerevisiae* Sir3 protein BAH (Bromo Associated Homology) domain loop 3. The BAH domain interacts with the Nucleosome core (Armache et al., 2011; Wang et al., 2013a) and BAH domain loop3, which showed structural and primary sequence conservation with Chp1CD LOOP1 (Figure 2.2 B), interacts with the Nucleosome core at the interface between histone H3, H2B and H4. The BAH domain Nucleosome binding mode is similar to what we observed for Chp1CD in our preliminary cryo-EM maps (data not shown). All mutants were expressed and purified in high amounts from the *E.coli* BL21 (DE3) pLysS expression strain (Figure 2.5 A). As illustrated in Figure 2.3 D and in Figure 2.4 D, in Nucleosome pull down assays, mutations in the Chp1CD loops already destabilized the binding with H3K9me3 Nucleosomes. The efficiency of Nucleosome pull down already diminished in the LOOP2B mutant and it was further impaired in the LOOP1A mutant and in the LOOP1A/LOOP2B and LOOP1B/2B combinations. A similar trend was observed for wild-type Nucleosome pull downs (Figure 2.4 E): in this case almost no Nucleosome was bound in LOOP1A/2B and LOOP1A/2A mutants, while LOOP1A and LOOP2B mutants were already showing a defective pull down efficiency.

Results

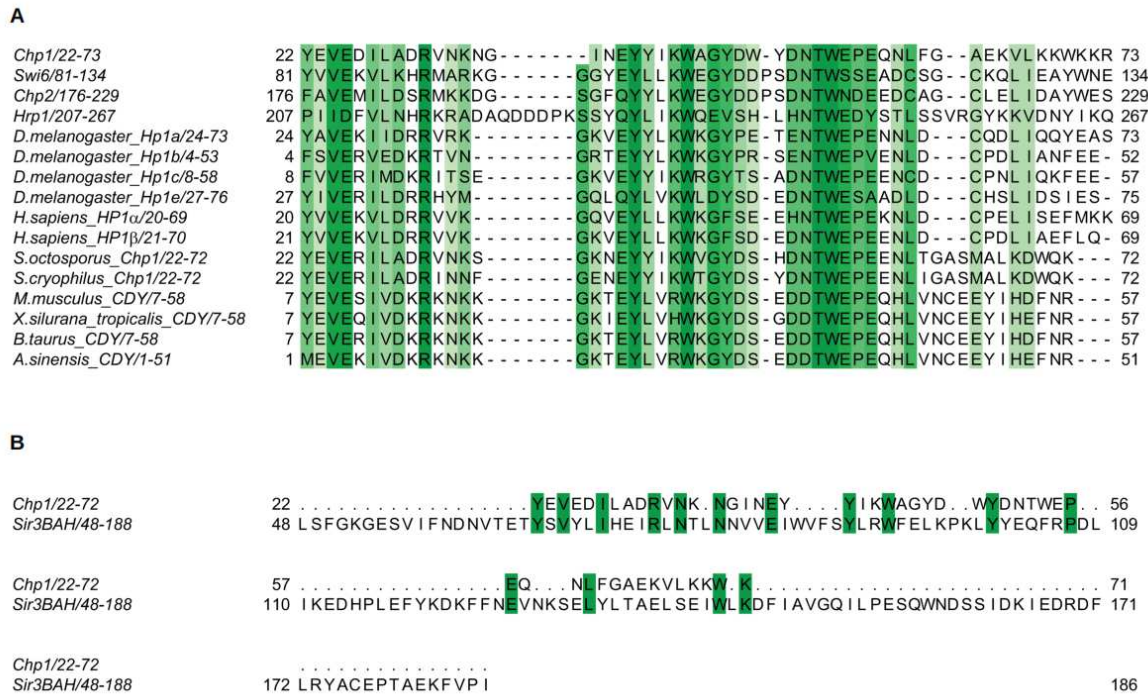


Figure 2.2: Sequence alignment of Chp1 chromodomain

(A) Sequence alignment of Chp1CD with chromodomains from fission yeast, *Drosophila melanogaster* and several mammals. The LOOP1/LOOP2 residues mutated show some conservation: R31, D51 and N52 are present in most chromodomains analyzed, whereas N33 is conserved only in a small group. The different shades of green correspond to a different degree of conservation (darker green is higher conservation). (B) Sequence alignment of the Sir3 protein BAH domain (from *S.cerevisiae*) with the Chp1CD. The residues in Chp1CD LOOP1 interacting with the Nucleosome core are conserved also in Sir3BAH.

Nucleosome without histone tails were generated (Tailless Nucleosomes, Figure 2.4 A) by Trypsin digestion, to determine whether the Chp1CD-H3K9me3 Nucleosome complex instability observed was determined by a decreased H3 tail binding affinity or by Nucleosome core interaction loss. The wild-type Chp1CD was able to bind to Nucleosomes in the absence of histone tails (Figures 2.3 C and 2.4 C). This evidence proved biochemically what previously observed in the preliminary cryo-EM structures: Chp1CD could bind to the Nucleosome core independently of the histone tails. However, given the amount of Nucleosome input applied in the pull down assays and comparing the various Nucleosome species, it could be concluded that the highest affinity is always determined by the presence of a methylated histone H3 tail, in accordance to the published literature (Schalch et al., 2009). In the cryo-EM density maps (data not shown) Chp1CD was contacting the Nucleosome core through protein-protein interactions, with the main contact between the Chp1CD LOOP1 and the histone H3 loop (residues 77-81).

Results

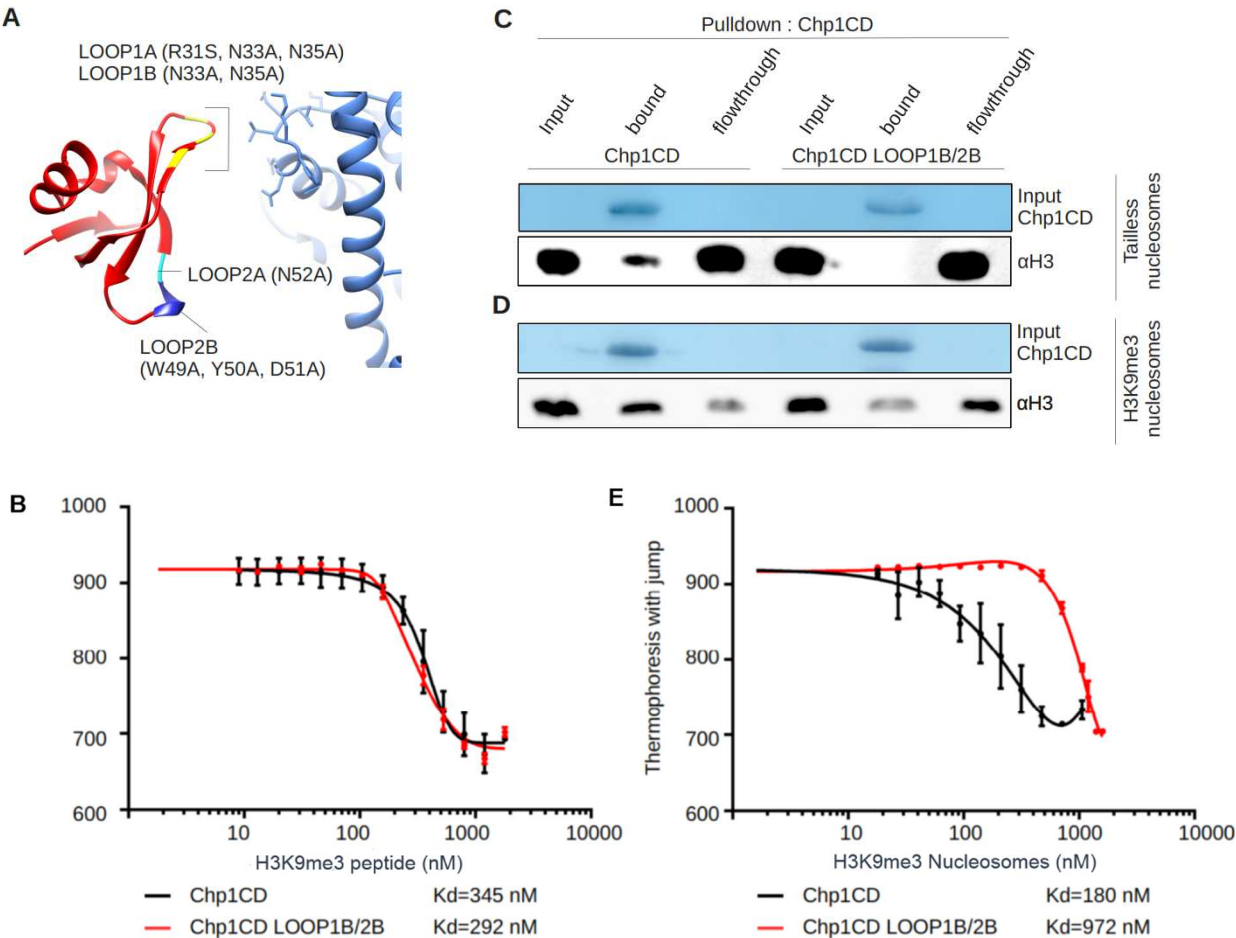


Figure 2.3: Mutations in Chp1CD LOOP1 and LOOP2 destabilize the interaction with the Nucleosome

A) Schematic representation of the mutations introduced on the Chp1CD loops (model produced by Prof. Dr. Mario Halic). Highlighted in yellow the three mutations on LOOP1 (R31S, N33A and N35A). In blue, the mutations introduced on LOOP2: in light blue the LOOP2A mutation (N52A) and in dark blue the LOOP2B mutations (W49A, Y50A and D51A). In blue on a side, the nucleosome is shown, with the side chain of the H3K79 residue protruding, close to the Chp1CD. (B) Microscale thermophoresis assay (MST) showing the binding affinity of the Chp1CD for an H3K9me3 peptide after introducing the LOOP1B/2B mutations, compared to the wild-type protein. (C) Tailless Nucleosome pull down assay showing that the interaction with the core of the Nucleosome is lost in the LOOP1B/2B mutant. (D) H3K9me3 Nucleosome pull down assay showing that the stability of the Chp1CD-H3K9me3 Nucleosome complex is diminished in the LOOP1B/2B mutant. (E) MST assay showing the binding affinity of the Chp1CD for an H3K9me3 Nucleosome after introducing the LOOP1B/2B mutations, compared to the wild-type protein.

To investigate the importance of this interaction for Chp1CD core binding, Chp1CD LOOP1 and LOOP2 mutants were tested for the ability to pull down Tailless Nucleosomes. In LOOP2B and LOOP1A mutants a several time reduction in the pulled-down Tailless Nucleosome signal was observed (Figure 2.4 C). This indicated that destabilizing one of the two Nucleosome core interacting loops was already sufficient to impair the affinity for Tailless Nucleosomes. A complete loss of binding was observed when destabilizing both interacting loops (LOOP1A/2B, LOOP1A/2A, LOOP1B/2B combinations (Figures 2.3 C and 2.4 C). The cumulative experimental evidence prompted us to conclude that the Nucleosome core

Results

interaction is dependent on the Chp1CD LOOP1 and LOOP2 integrity. As already mentioned a decreased affinity was observed in LOOP1/2 mutants for the H3K9me3 Nucleosome, which is the natural binding partner of the Chp1CD *in vivo* (Figures 2.3 D and 2.4 D). To determine the contribution of the Nucleosome core binding to the stability of the Chp1CD-H3K9me3 Nucleosome complex, the affinity of the LOOP1 and LOOP2 mutants for the H3K9me3 tail was measured. This allowed to discriminate between the histone H3 tail and the Nucleosome core contributions to the stability of the Chp1CD-H3K9me3 Nucleosome complex.

Therefore, H3K9me3 peptide pull down and equilibrium-in solution H3K9me3 peptide binding assays (Microscale Thermophoresis) with LOOP1 and LOOP2 mutants were set up to quantify the binding affinity for the H3 methylated tail.

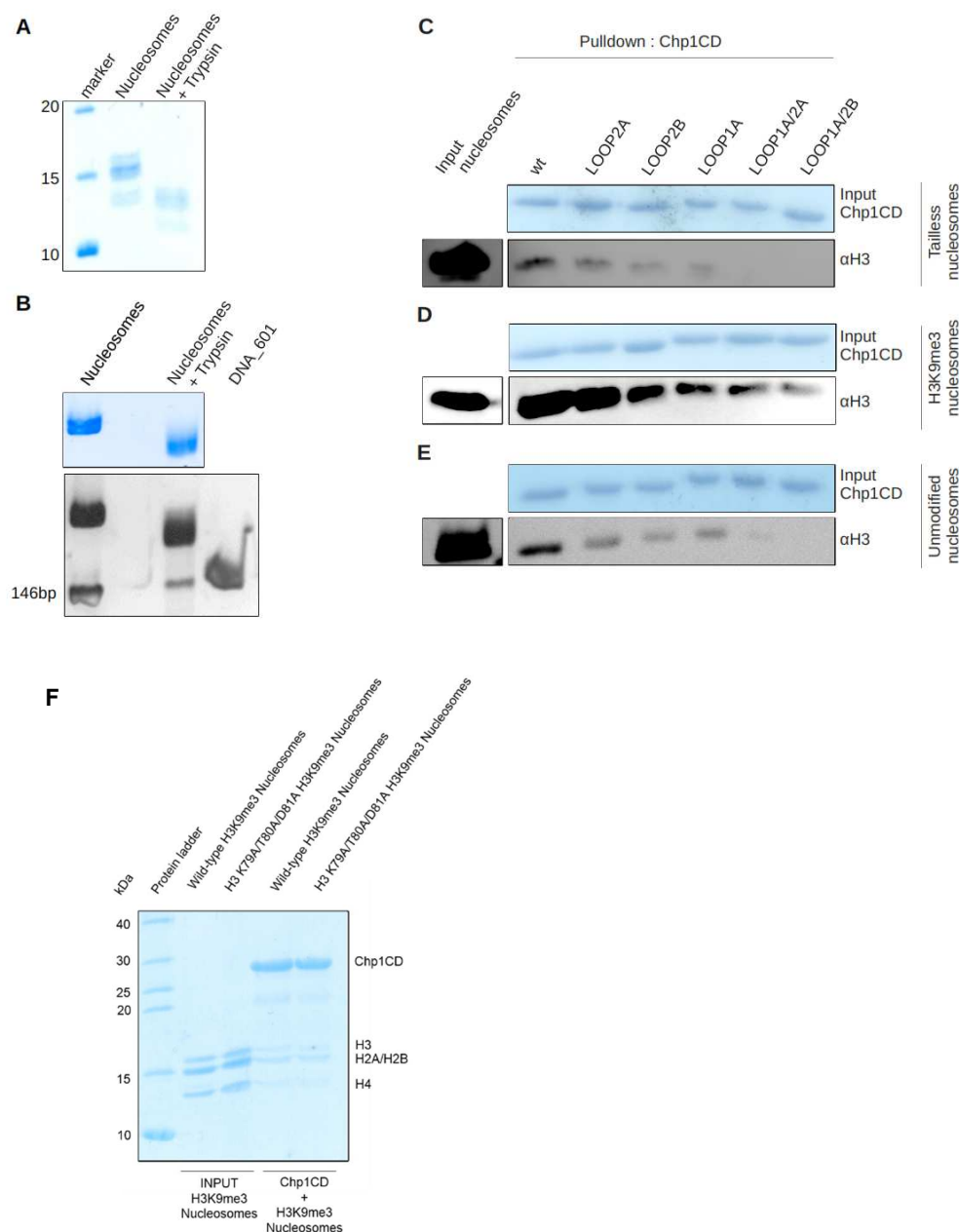


Figure 2.4: Chp1CD-Nucleosome interaction is destabilized in LOOP1/2 mutants (in the previous page)

A) Coomassie Blue stained 15% acrylamide gel showing the Nucleosome digestion by Trypsin, with a shift in molecular weight for each histone band in the treated sample. (B) Native-TBE acrylamide gel run of Nucleosome and Nucleosome trypsinized (Tailless) stained with Coomassie Blue and SyBrGOLD. Tailless Nucleosomes are stable after Trypsin digestion. (C) Tailless Nucleosome pull down assay showing that the interaction with the core of the Nucleosome is lost in the LOOP1A/2A and LOOP1A/2B mutants. (D) H3K9me3 Nucleosome pull down assay showing that the interaction with the Nucleosome is impaired in the LOOP1/2 mutants. (E) Unmodified Nucleosome pull down assay showing that the interaction with the core of the Nucleosome is lost in the LOOP1A/2A and LOOP1A/2B mutants. (F) Coomassie Blue stained 15% acrylamide gel of an H3K9me3 Nucleosome pull down assay. Chp1CD was bound to a Ni-NTA resin and incubated with two types of H3K9me3 Nucleosomes: the wild-type H3K9me3 Nucleosome and a mutated version, harboring the H3K79A/T80A/D81A mutations of the Chp1CD interacting H3 loop. As shown, Chp1CD binding to H3K9me3 Nucleosomes is not destabilized by the presence of additional mutations on the histone H3 globular domain, most likely because the interaction interface with the Nucleosome core is degenerated, with multiple contact points. A similar result was obtained with the Sir3 BAH domain, unaffected by single mutations on the Nucleosome interacting interface (Wang et al., 2013a).

In the H3K9me3 peptide pull down assays, all Chp1CD mutants retained the ability of binding the H3K9me3 histone peptide as the wild-type protein (Figure 2.5 B and C). As comparison for complete loss of binding, the previously characterized W44A mutant, affecting the stability of the Chp1CD aromatic cage (Ishida et al., 2012), was used as negative control. Mutating residues N33A and N35A in LOOP1 did not affect histone H3 peptide-Chp1CD binding. The same occurred in LOOP2B (three mutations: W49A/Y50A/D51A) and LOOP2A (one mutation in the conserved N52 residue: N52A) mutants. One residue revealed to be important for the H3K9me3 peptide binding stability, namely Chp1CD R31. In all the mutants harboring the R31S mutation (R31S single mutant, LOOP1A, LOOP1A/2A and LOOP1A/2B), the Chp1CD binding affinity dropped to roughly 50% of the wild-type protein in the pull down experiments (Figure 2.5 C). Therefore, in this set of mutants (LOOP1A, LOOP1A/2A, LOOP1A/2B) it was not possible to discriminate if the lack of H3K9me3 Nucleosome binding stability observed (Figure 2.4 D) was due to a loss of affinity for the H3K9me3 histone tail (Figure 2.5 C), for the lack of binding to the Nucleosome core (Figure 2.4 C) or the cumulative effect of the two. On the other hand, mutants LOOP1B (N33A and N35A mutated) and LOOP1B/2B (N33A, N35A and W49A/Y50A/D51A mutated) were generated by excluding the R31A mutation from LOOP1A and LOOP1A/2B mutants. These mutants showed wild-type levels of binding for the H3K9me3 peptide in the pull down assays (Figure 2.5 B and C). To confirm the pull down experiments on LOOP1B and LOOP2B mutants another method was used in parallel, the Microscale Thermophoresis (MST).

Results

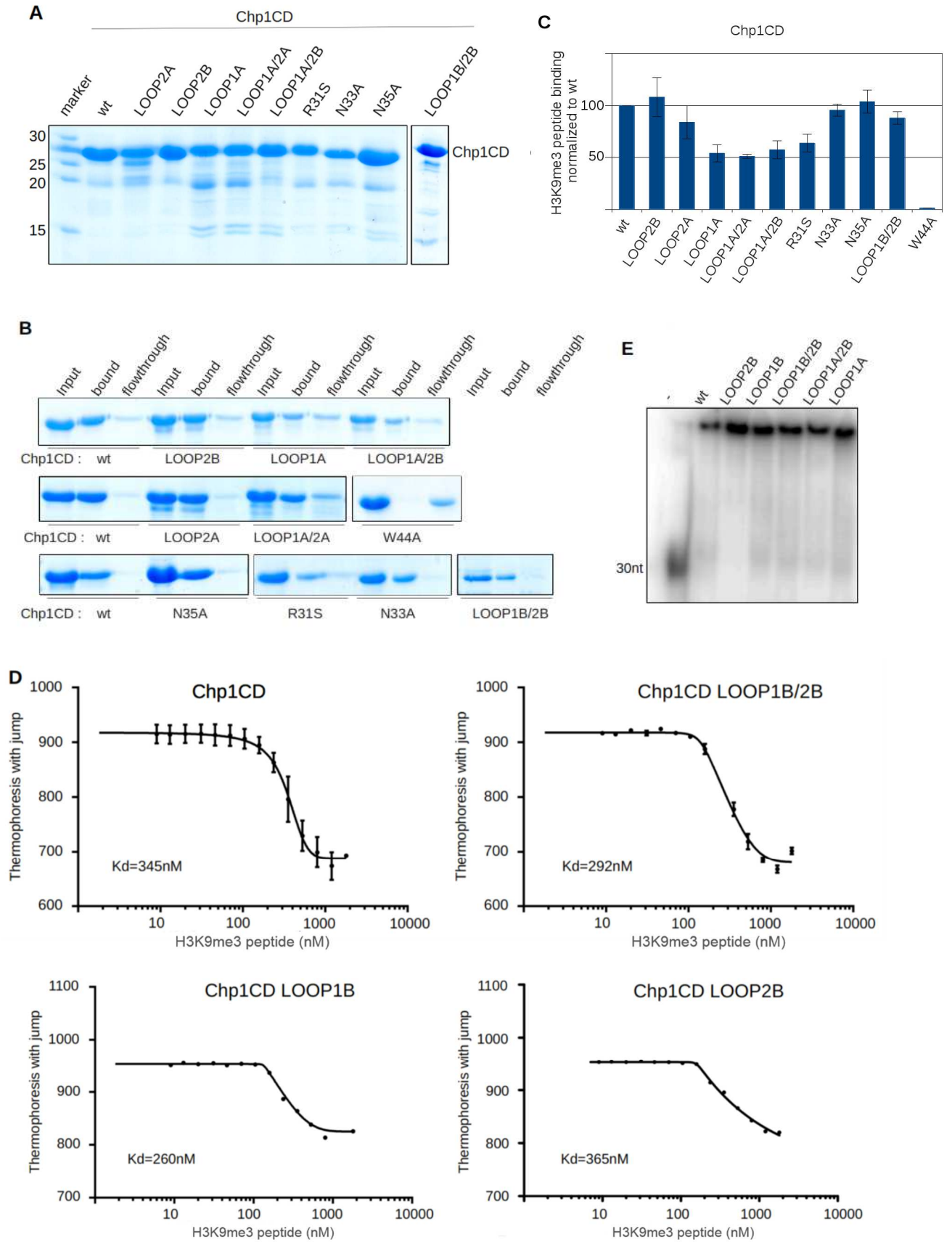


Figure 2.5: LOOP1B/2B mutations do not affect the binding affinity for the H3K9me3 peptide or the RNA (in the previous page)

A) Coomassie stained SDS-PAGE showing the purified Chp1CD wild-type and mutant proteins. (B) Peptide pull down assay with various Chp1CD mutants. The binding affinity is normalized to the wild-type Chp1CD in each assay. (C) Quantification of the H3K9me3 peptide pull down shown in (B). 3 independent repetitions were averaged. The Chp1CD mutants bearing the R31S mutation in LOOP1 showed a prominent defect in H3K9me peptide binding. LOOP2A, LOOP2B, N33A, N35A and LOOP1B/2B mutants show no significant difference compared to the wild-type protein. W44A mutant was used as a control for no binding (Ishida et al., 2012). (D) Microscale thermophoresis H3K9me3 peptide binding assay on the wild-type Chp1CD and the LOOP1B, LOOP2B, LOOP1B/2B mutants. Calculated K_d is shown below each binding curve. (E) RNA shift assay on a TBE-Native gel showing that LOOP1 and LOOP2 mutants bind centromeric RNAs as the wild-type Chp1CD (performed by Dr. Mirela Marasovic).

MST is commonly used to measure the binding affinity of proteins for their interaction partners in solution, by estimating their ability to move under infrared beam heat: if a protein is interacting in a complex, its thermal motion in solution would be slower compared to the protein alone. This can be measured by fluorescently labeling the protein of interest and measuring the local decrease in fluorescence over the time while the infrared stimulus is applied. By increasing the amount of one of the interacting proteins while keeping constant the concentration of the fluorescently labeled one, it is possible to obtain a binding affinity curve and estimate the dissociation constant (K_d). Therefore, the Chp1CD wild-type protein, the LOOP1B, LOOP2B and LOOP1B/2B mutants were fluorescently labeled and two binding experiments were set up: one with a crescent concentration of H3K9me3 peptide and one with a crescent concentration of H3K9me3 Nucleosome. We measured the dissociation constants for each H3K9me3 peptide binding assay and we did not observe substantial differences in affinity for the Chp1CD wild-type ($K_d=345$ nM) and the LOOP1B ($K_d=260$ nM), the LOOP2B ($K_d=365$ nM) and the LOOP1B/2B ($K_d=292$ nM) mutants (Figure 2.3 B and 2.5 D). These results confirmed what previously emerged from the H3K9me3 peptide pull down assays, showing that mutations in LOOP1 and LOOP2 had no effect on H3K9me3 tail affinity unless residue LOOP1 R31 was modified. On the other hand, LOOP1B/2B mutations had huge impact on the stability of the Chp1CD-H3K9me3 Nucleosome complex, with a drop of about five fold in the affinity for H3K9me3 Nucleosomes from the wild-type protein ($K_d=180$ nM) to the LOOP1B/2B mutant ($K_d=972$ nM) (Figure 2.3 E). Despite the binding affinity for the H3K9me3 tail remained unchanged (Figure 2.3 B and 2.5 D), both the pull down assays (Figure 2.3 D) and the MST (Figure 2.3 E) showed that LOOP1B/2B mutations cause a drop in affinity for the H3K9me3 Nucleosome which is independent from the presence of the H3K9

methylation mark. The loss of interaction with the Nucleosome core observed for all LOOP1/2 mutant combinations (particularly in LOOP1A, LOOP1A/2A, LOOP1A/2B, LOOP1B/2B - Figures 2.3 C and 2.4 C) explains the importance of the integrity of Chp1CD LOOP1 and LOOP2 for the Nucleosome core binding and the stability of the Chp1CD-H3K9me3 Nucleosome complex. In particular, the specific protein-protein interaction occurring between the Chp1CD loops (LOOP1 and LOOP2) and histone globular domains of the Nucleosome core seems necessary to maintain the stability of the Chp1CD-H3K9me3 Nucleosome complex. However, comparing the strength of pull down of the wild-type Chp1CD for the H3K9me3 Nucleosomes, Unmodified Nucleosomes and Tailless Nucleosomes, it remained clear that the main recruitment of Chp1CD occurred only through the deposition of the H3K9 methylation mark (compare intensities of input and Pulled-down Nucleosomes for the wild-type Chp1CD in Figure 2.3 C-D and Figure 2.4 C-D-E). The binding to the Nucleosome core could contribute to give to Chp1 higher affinity for the Nucleosome target (only after H3K9me3 tail is bound) and a less dynamic behavior, to allow for timely recruitment of the RITS-Argonaute to establish heterochromatin. These biochemical evidences correlated very well with the structural model proposed and induced us to better investigate the importance of the Chp1CD-Nucleosome core interaction in heterochromatin maintenance and establishment.

Chp1CD binding to the Nucleosome core is required for heterochromatin formation

To understand the importance of core interaction in Chp1 function, the ability of our Chp1 mutants to rescue defective heterochromatin *in vivo* was tested. The *chp1*⁺ wild-type gene and the *chp1* LOOP1/2 mutated versions (the entire *chp1*⁺ transcriptional unit) were cloned on a pREP1 plasmid under the *chp1*⁺ gene endogenous promoter (replacing the original pREP1 *nmt1*⁺ promoter) and transformed in a *chp1Δ* strain. Growth assays under selective FOA medium showed that when transforming the *chp1Δ* strain with the wild-type *chp1*⁺ gene a full silencing rescue was obtained (Figure 2.6, *chp1Δ* + *chp1*, compared to wild type). Similarly, all those chromodomain mutants that in the *in vitro* assays (pull downs and MST) did not show any affinity defects for the H3K9me3 peptide or the H3K9me3 Nucleosome (Figure 2.3, 2.4 and 2.5), did not have any growth defect under selective FOA medium and were able to fully restore heterochromatin, rescuing the *chp1Δ* strain phenotype (Figure 2.6 A and B, LOOP2A, LOOP2B, LOOP1B).

Results

By contrary, transformation of LOOP1A, LOOP1A/2A, LOOP1A/2B, LOOP1B/2B mutants did not lead to the rescue of the *chp1Δ* strain centromeric silencing defect (Figure 2.6). These mutants were previously observed having a defect in binding H3K9me3 Nucleosomes (Figures 2.3 D and 2.4 D) and Tailless Nucleosomes (Figures 2.3 C and 2.4 C) in pull down assays.

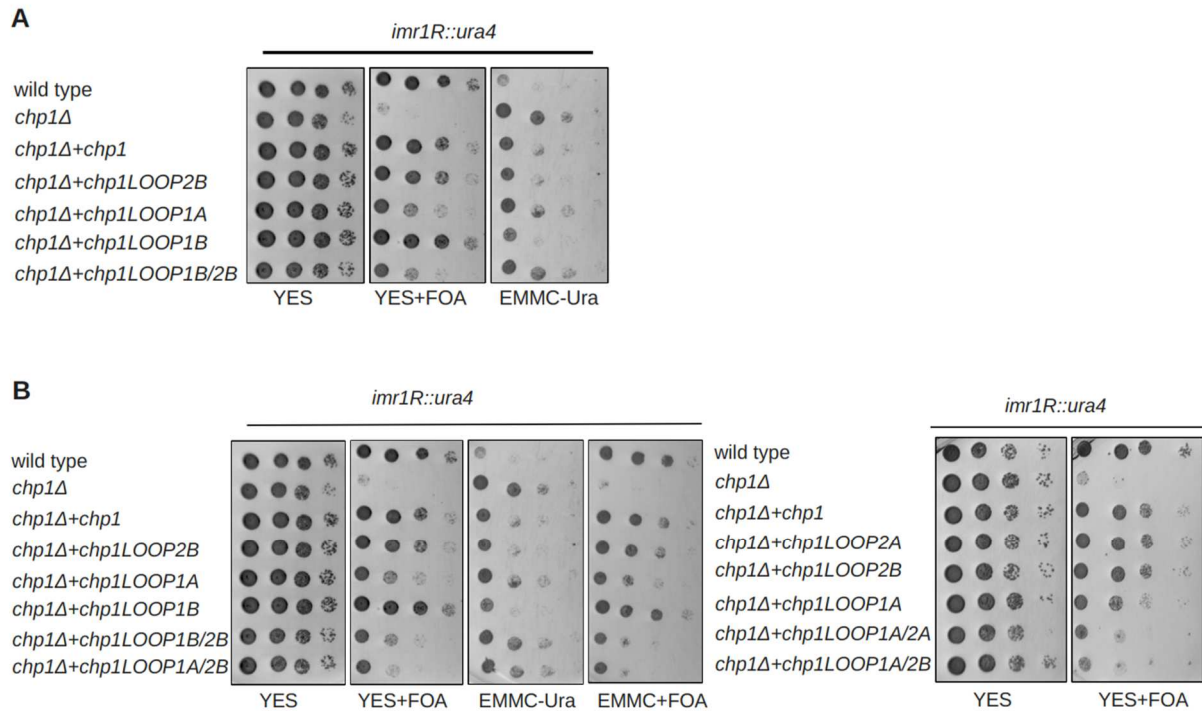


Figure 2.6: Chp1CD interaction with the Nucleosome core is required for heterochromatin formation

(A-B) Silencing assays showing the effect of LOOP1/2 mutations on centromeric heterochromatin. 10-fold serial dilutions were spotted on YES (non-selective), YES+FOA, EMMc+FOA, EMMc-Ura media.

These evidences highlight the importance of the stability of the Nucleosome core interaction for heterochromatin silencing. The partial defective phenotype of the LOOP1A mutants (Figure 2.6 A-B) could be derived from a combination of loss of affinity for both the H3K9me3 histone tail (Figure 2.5 C) and the Nucleosome core (Figure 2.4 C). By combining LOOP1 and LOOP2 mutations, the centromeric silencing was further impaired (Figure 2.6 A-B, LOOP1A/2A, LOOP1A/2B, LOOP1B/2B mutants) correlating with the need for both LOOP1 and LOOP2 in Chp1CD interaction with the Nucleosomal core interface.

Results

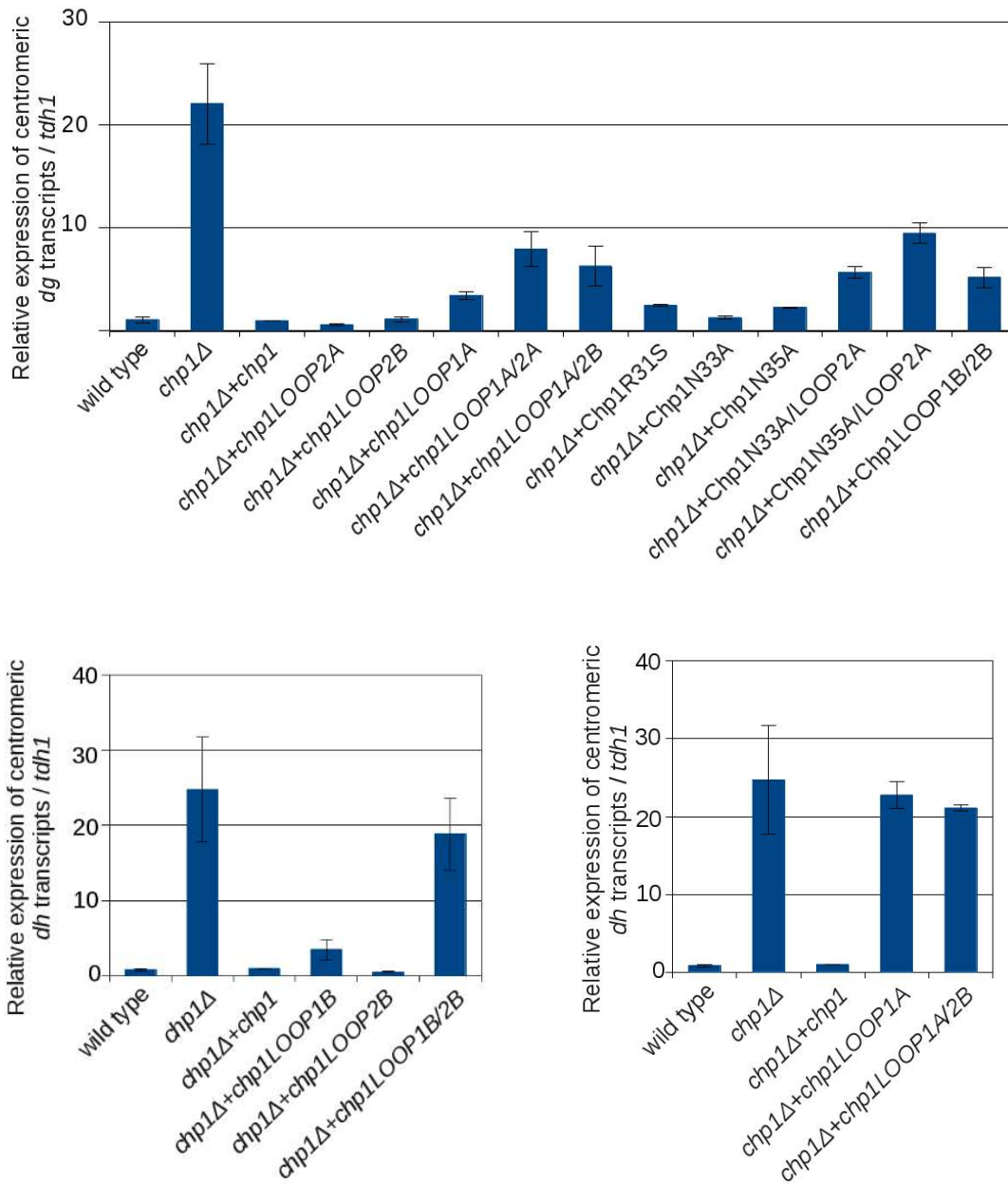


Figure 2.7: Relative expression of dg and dh centromeric transcripts in wild-type and Chp1CD mutants

RT-qPCR experiments quantifying the centromeric RNA levels in wild-type and Chp1CD mutant cells. LOOP1/2 mutants showed the highest enrichment over wild-type both at dg (above) and dh repeats (bottom), correlating with the silencing defect observed in the growth assays. dh repeats showed a more prominent defect in silencing, with LOOP1/2 mutants having a dh RNA upregulation comparable to the chp1 deletion strain. Error bars indicate the Standard Error Mean (SEM) of three independent experiments.

Results

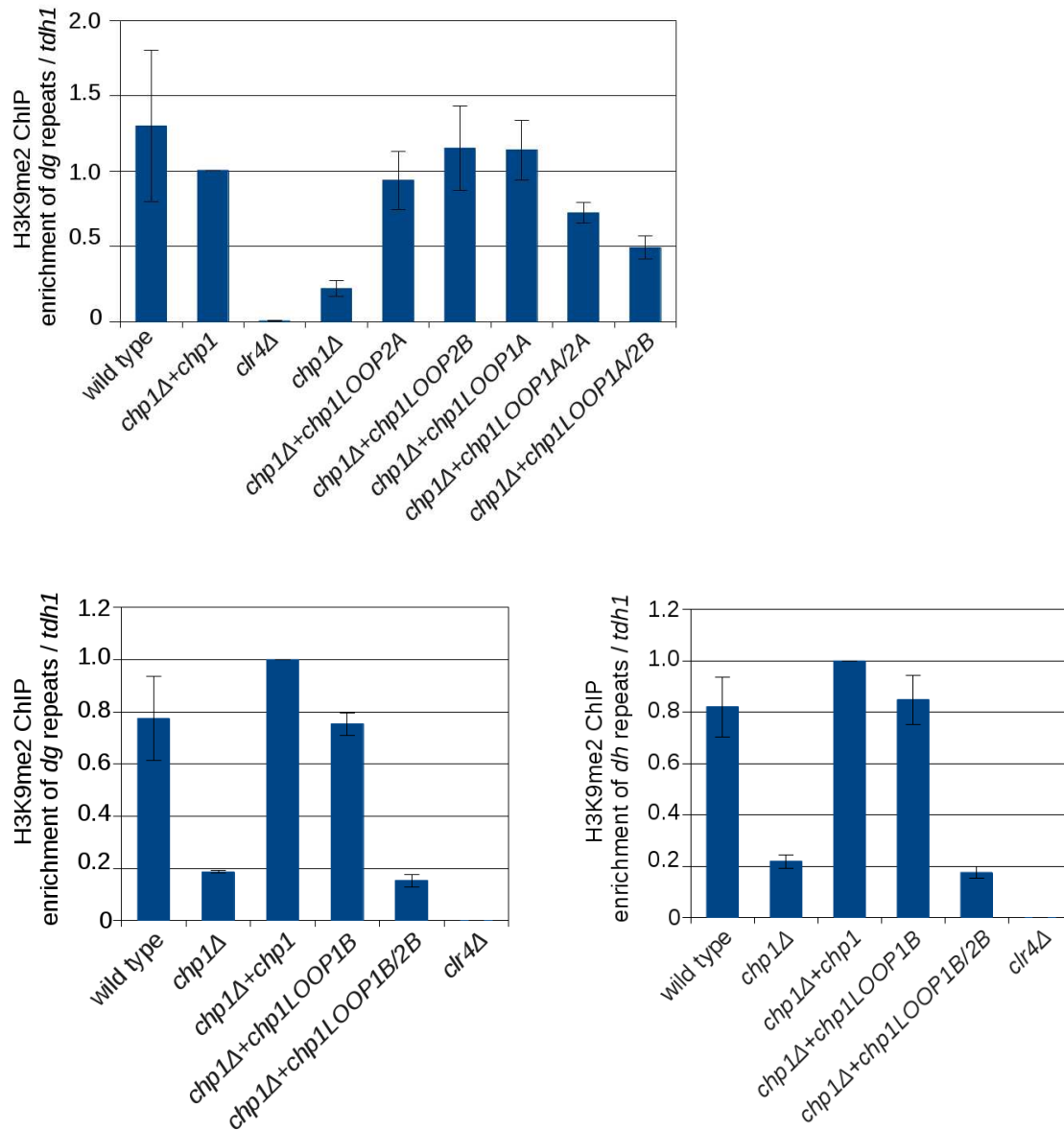


Figure 2.8: Centromeric H3K9me2 levels are decreased in Chp1CD LOOP1/2 mutants

ChIP-qPCR experiments measuring the levels of H3K9me2 at centromeric dg and dh repeats. LOOP1/2 mutants showed a prominent H3K9me2 level drop at both centromeric dg and dh repeats. Error bars represent the Standard Error Mean (SEM) of three independent experiments.

To understand more in detail the contribution of Nucleosome core binding to the described phenotype, the attention was focused on the LOOP1B/2B mutant, which did not present any affinity change for the H3K9me3 histone tail compared to the wild type Chp1CD (Figure 2.3 B). As shown previously, the affinity for the H3K9me3 Nucleosome *in vitro* sensibly drops

Results

(almost 5 fold, MST, Figure 2.3 E), despite LOOP1B/2B still retained the ability to bind the H3 methylated tail. This was due to a lack of interaction with the Nucleosome core (Figure 2.3 C) and likely explained why the phenotype observed *in vivo* for LOOP1B/2B was not as severe as for LOOP1A/2A and LOOP1A/2B mutants, which had, on top of impaired core binding, also a lower affinity for H3K9me3 tail (Figure 2.5 C). To further confirm the phenotype screening, the centromeric RNA expression levels were measured by using RT-qPCR. LOOP1/2 mutant combinations had the strongest silencing defects, both at centromeric *dg* and *dh* repeats (Figure 2.7). Already an N35A single mutations in LOOP1 caused a slight upregulation in the centromeric *dg* transcripts. LOOP1A had a fivefold increase in *dg* RNA levels whereas the combination of LOOP1A with LOOP2 mutations had almost 10 fold increase.

Together with centromeric RNA silencing, the levels of H3K9me2 deposition at centromeres were monitored, to verify whether LOOP1/2 Chp1CD mutants had lower levels of the histone mark compared to wild-type cells.

As envisioned, the LOOP1/2 mutants (LOOP1A/2A, LOOP1A/2B and LOOP1B/2B) had the strongest loss of H3K9me2 at centromeres, further confirming the results obtained with the silencing assays and the RT-qPCR experiments for the centromeric RNA expression (Figure 2.8). Moreover, ChIP experiments with an antibody specific for the Chp1 protein showed that the defect in silencing at centromeres is accompanied by a loss of Chp1 protein localization both at *dg* and *dh* repeats (Figure 2.9). In fact, LOOP1A/2A and LOOP1A/2B had a full defect in localizing Chp1 at the centromeric region (Figure 2.9). LOOP1B did not present any major impairments, while LOOP1B/2B had also a prominently impaired Chp1 localization (Figure 2.9). The loss of localization indicated that the Nucleosome core interaction is indeed affecting the stability of Chp1 on chromatin. Since the LOOP1B/2B mutant binds the histone H3 tail as the wild-type Chp1CD, the recruitment disability could arise only from a defective core binding. Moreover, published results (Schalch et al., 2009) underlined that a drop in affinity of the Chp1CD for the H3K9me3 histone tail in the range of the one we observed for the Nucleosome (5-fold for the H3K9me3 Nucleosome, Figure 2.3 E) did not have any major repercussions *in vivo* on heterochromatin maintenance. This further proves that Chp1CD Nucleosome core interaction is indeed required for a specialized function which is beyond histone H3 tail recognition. Since the observed phenotypes could have been caused by a difference in expression of the mutant Chp1 proteins, the expression levels of the Chp1 protein in our heterochromatin rescue mutants were quantified by Western blot and compared to the wild-type rescue. All the Chp1 mutant proteins are expressed at similar levels as the wild-type

protein, further validating the results obtained (Figure 2.10).

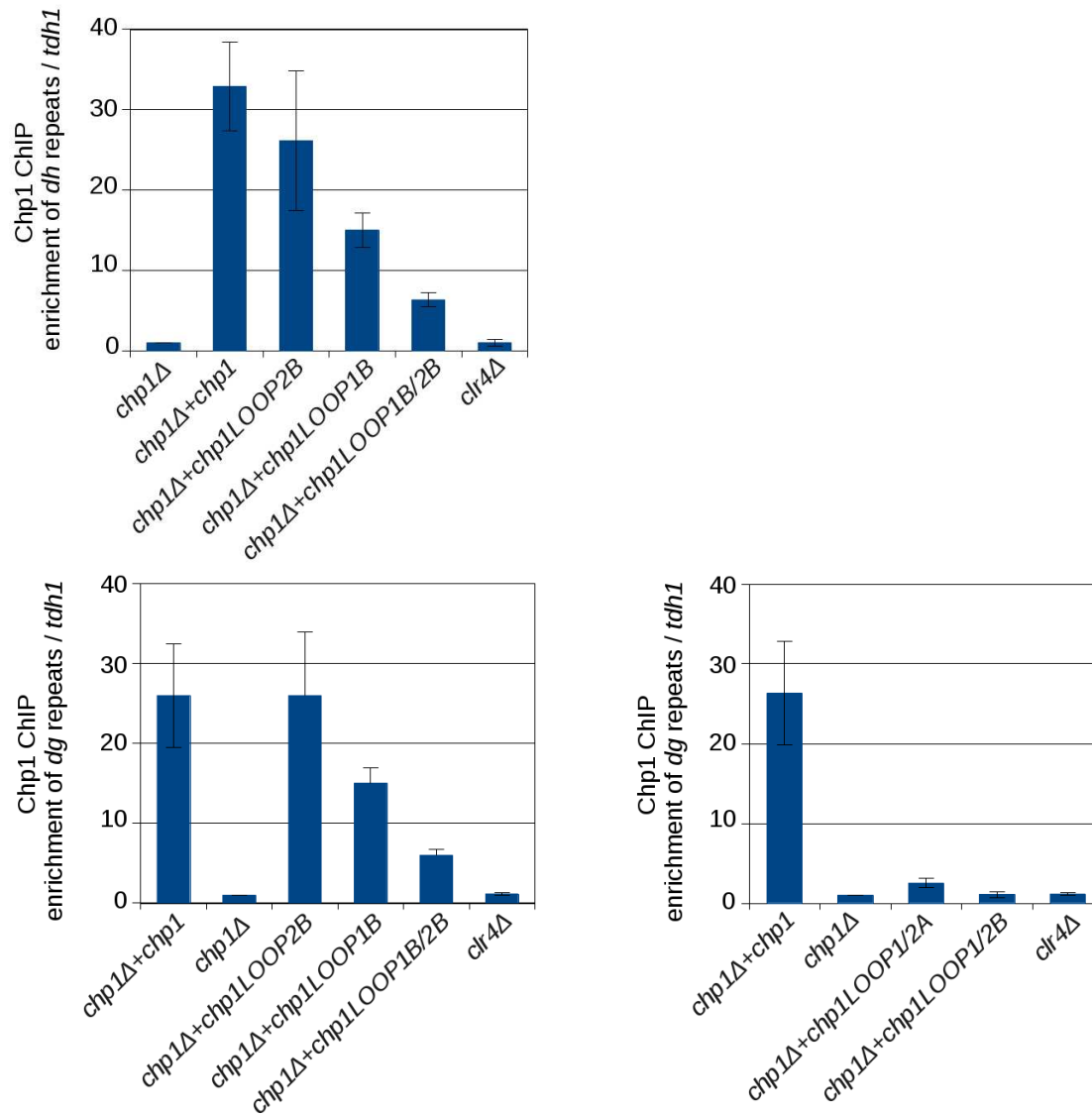


Figure 2.9: Chp1 localization at centromeres is compromised in *LOOP1/2* mutants

Chp1 ChIP experiments showing the loss of localization of the *Chp1* protein when *LOOP1/2* mutations are present. Error bars represent the Standard Error Mean (SEM) of three independent experiments.

Previously published results (Ishida et al., 2012) showed the importance of the positively charged residues in the C-terminal α -helix of the Chp1CD in binding either DNA or RNA *in vitro* and centromeric RNAs *in vivo*. Chp1CD α -helix mutants that were not able to bind nucleic acids *in vitro* had also a defect in heterochromatin formation.

Therefore Chp1CD mutants were tested for centromeric RNA binding affinity. In fact, *LOOP* mutations could possibly have an indirect effect on RNA binding that would explain the observed *in vivo* loss of silencing. Our Chp1CD mutants however bind the centromeric RNA

(30 nt centromeric RNA) as the wild-type protein in *in vitro* RNA native EMSA assays (Figure 2.5 E).

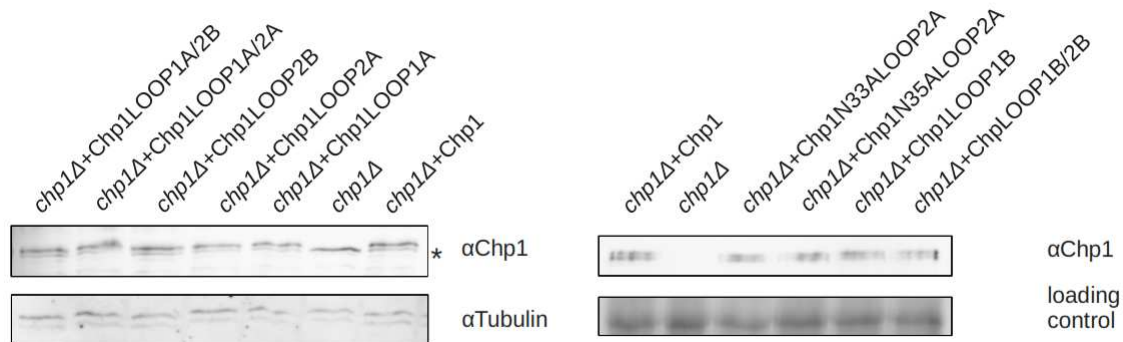


Figure 2.10: The Chp1 proteins (wild-type and Chp1CD mutants) are expressed at the same levels *in vivo*

Western-blot experiment on whole cell lysates with a Chp1 specific antibody. As loading controls, we used either an anti-Tubulin antibody or the Amido-black stained membrane. * (on the left panel) marks a non-specific band present in all the tested cell extracts.

This directed to the conclusion that the observed heterochromatin loss was due to an impaired Chp1CD-Nucleosome core interaction. In conclusion, the overall *in vivo* evidences presented in this sub-chapter highlighted the importance of Chp1CD Nucleosome core interaction for Chp1 proper localization and function in heterochromatin formation.

Genomic integration of the LOOP1B/2B mutant causes loss of heterochromatin silencing at centromeres

The Chp1 heterochromatin rescue experiments showed that mutations in LOOP1/2 are effective in determining centromeric loss of silencing. It was still under question whether the phenotype observed might have been induced from the slight over-expression of the rescuing Chp1 protein (expressed under its endogenous promoter on a pREP1 plasmid modified to remove the *nmt1*+ promoter), derived from the ectopic plasmid being in multiple copies in the cell. The wild-type Chp1 protein was completely rescuing the *chp1Δ* strain heterochromatin defect, bringing the silencing back to the *chp1*+ wild-type strain (Figure 2.6, 2.7 and 2.8- compare the rescue *chp1Δ*+*chp1* with the wild-type strain). Therefore, the minimal over-expression of Chp1 was not toxic for cell growth or heterochromatin assembly. Moreover, all the Chp1 proteins were expressed (wild-type and mutants) at similar levels in the rescue strains

Results

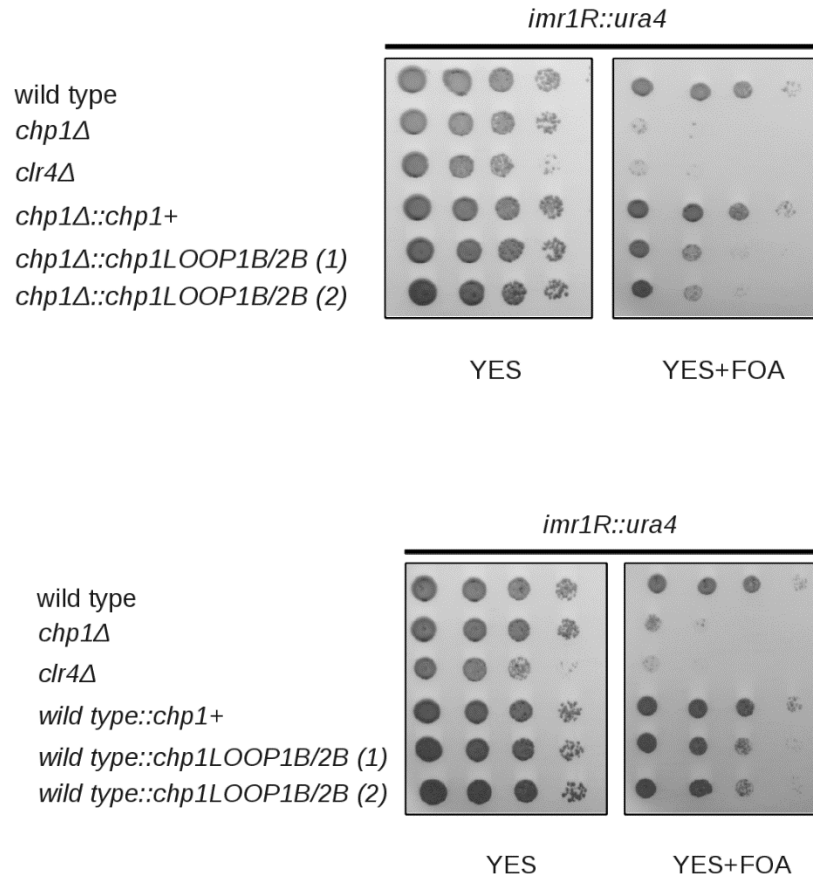


Figure 2.11: The genomic integration of LOOP1B/2B mutations causes a heterochromatin defective phenotype

Silencing assays showing that the genomic insertion of LOOP1B/2B mutations in a wild-type (SPY28) background –panel above- or in a *chp1Δ* (SPY28) background –panel below- causes a defect in heterochromatin maintenance comparable to the one previously observed with the plasmid rescue experiments. 10-fold serial dilutions were spotted on YES and YES+FOA. Results are shown for two colonies of the LOOP1B/2B integration for both backgrounds, named LOOP1B/2B (1) and (2).

(Figure 2.10), indicating not only their stability, but also that the differences observed in the growth and molecular assays were arising from their intrinsic properties.

Nonetheless, as additional control, the Chp1LOOP1B/2B mutations were integrated in the genome, creating the following molecular construct:

(SphI) Region at the 5' of Chp1 gene (Chromosome I, 2215500 – 2215055) (AscI) – *HphMX6* resistance cassette – (SphI) Chp1 endogenous promoter (Chromosome I, 2214829 – 2214664) – Chp1 coding sequence – (BamHI) Chp1 terminator (Chromosome I, 2210976 – 2210582) (BamHI).

The wild type Chp1 and the Chp1LOOP1B/2B constructs were generated from the initial pREP1 plasmid by traditional cloning, amplified by PCR and transformed into both the wild-

Results

type and *chp1Δ* background strains (SP28 and SP170, genotypes reported in Materials and Methods, Table 3). The wild-type construct transformation was an important control to verify that the integration upstream of the *chp1+* gene of the hygromycin antibiotic resistance cassette (*hphmx6*) into the wild-type or the *chp1Δ* strain would not cause a heterochromatin defect by itself. The genomic integration of the wild-type Chp1 did not have any phenotypic effect under selective medium (YES+FOA), indicating that the protein is functional in both the wild-type and the *chp1Δ* backgrounds (Figure 2.11). However, Chp1LOOP1B/2B genomic introduction caused a heterochromatin defective phenotype, totally similar to what observed in the plasmid rescue assays (Figure 2.6, compare with the Chp1LOOP1B/2B in Figure 2.11). The result was confirmed by checking the expression levels at *dg* and *dh* centromeric repeats (Figure 2.12). In both backgrounds, an upregulation of centromeric transcripts (> 20 fold) at *dg* and *dh* repeats for the Chp1LOOP1B/2B mutant was observed, with a slightly higher defect at *dh* repeats and overall higher defect than the one observed in the plasmid rescue experiments (Figure 2.7).

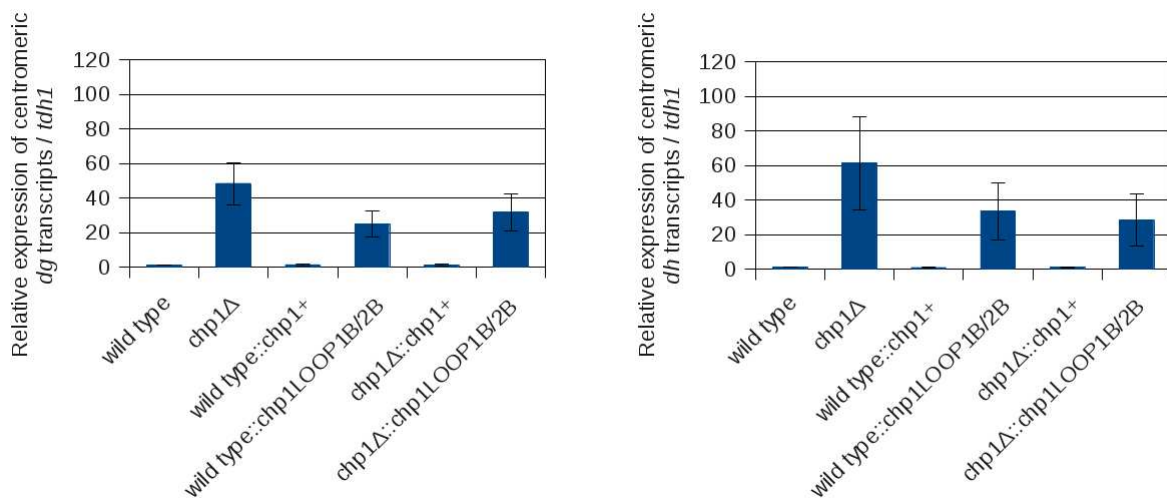


Figure 2.12: Centromeric RNA expression levels in wild-type and Chp1LOOP1B/2B genomic integration strains

Relative expression of centromeric dg (on the left) and dh transcripts (on the right) for the wild-type and Chp1LOOP1B/2B genomic integration strain. As controls, the wild-type and chp1Δ background strains were compared to the integration strains. Error bars represent the SEM on three independent RNA sets

There was not substantial difference regarding the two backgrounds tested: Chp1LOOP1B/2B integrated in both wild-type and the *chp1Δ* deletion strain caused similar silencing impairment at centromeres with upregulation of centromeric transcripts to similar values (Figures 2.11 and 2.12). The integrated wild-type Chp1 construct, on the other hand, was able to maintain

Results

heterochromatin to wild-type levels with no difference if compared to the wild-type background strain. The genomic integration of the LOOP1B/2B mutations gave us an important confirmation on how Nucleosome core binding is important for heterochromatin maintenance *in vivo*. Moreover, this added additional proof that no artifact phenotype was introduced by using the pREP1 modified Chp1 plasmid to rescue the *chp1Δ* defect (Figure 2.6, 2.7, 2.8 and 2.9), since the slight over-expression of Chp1 did not cause any heterochromatin defect.

Chp1CD-RNA binding stability is influenced by the Nucleosome core interaction

As already shown (Ishida et al., 2012), Chp1CD was able to bind nucleic acids and Chp1CD mutants in the residues responsible for this interaction had a defective heterochromatin maintenance. In this study, the observed Chp1CD mutant phenotypes could have been originated from a defect in RNA binding rather than in the Nucleosome core interaction. To address this question, Native RNA EMSA assays with two different length of centromeric *dg* RNA, namely 30 nt (Figure 2.5 E) and 100 nt (Figure 2.13) were performed.

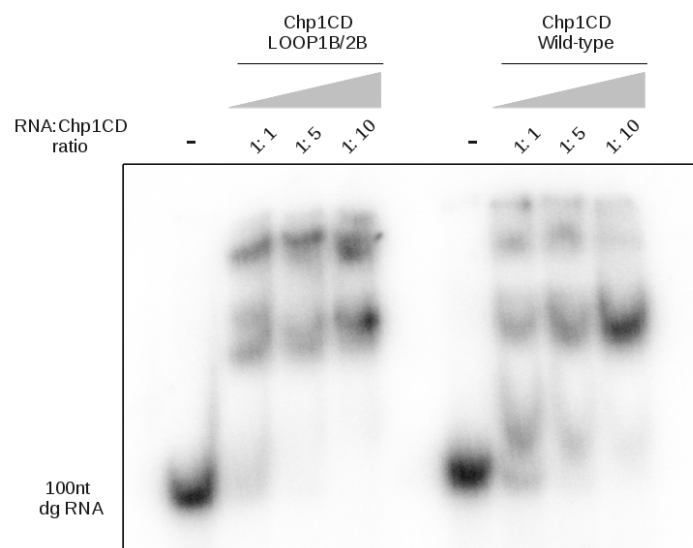


Figure 2.13: Chp1CDLOOP1B/2B binds to RNA with a slightly higher affinity than the wild-type Chp1CD

Gel shift assay with crescent amounts of Chp1CD (LOOP1B/2B and wild-type) incubated with ^{32}P radiolabeled 100 nt dg centromeric RNA and H3K9me3 peptide (1:1 Chp1CD:H3K9me3 peptide ratio). LOOP1B/2B binds to RNA with a slightly higher affinity compared to wild-type Chp1CD.

All the LOOP1/2 Chp1CD mutants analyzed for RNA binding still maintained the interaction (Figure 2.5 E). Particularly, the LOOP1B/2B mutant was interacting with the RNA with a

Results

slightly stronger affinity compared to the wild-type Chp1CD (Figure 2.13). All the binding assays were performed in the presence of H3K9me3 peptide, since LOOP2 mutations were already shown to have a defect when binding the RNA without the presence of the methylated H3 histone tail (Ishida et al., 2012).

It seemed therefore that the defect determined *in vivo* was not originating from a lack of affinity of our Chp1CD mutants for the RNA. Nonetheless, an RNA pull down assay was performed to test whether the lack of Nucleosome core interaction in the LOOP1B/2B mutant could indeed impair the efficiency of binding for centromeric RNA (Figure 2.14). ³²P radiolabeled 100nt *dg* centromeric RNA was incubated with Ni-NTA bound Chp1CD chromodomain (wild-type and LOOP1B/2B mutant) in the presence of H3K9me3 Nucleosomes (in a 1:1 Chp1CD: H3K9me3 Nucleosome ratio). Chp1CD LOOP1B/2B, despite binding the RNA with higher affinity when saturated with H3K9me3 peptide (Figure 2.13), is less able compared to the wild-type Chp1CD to pull down the RNA in the presence of H3K9me3 Nucleosomes (Figure 2.14).

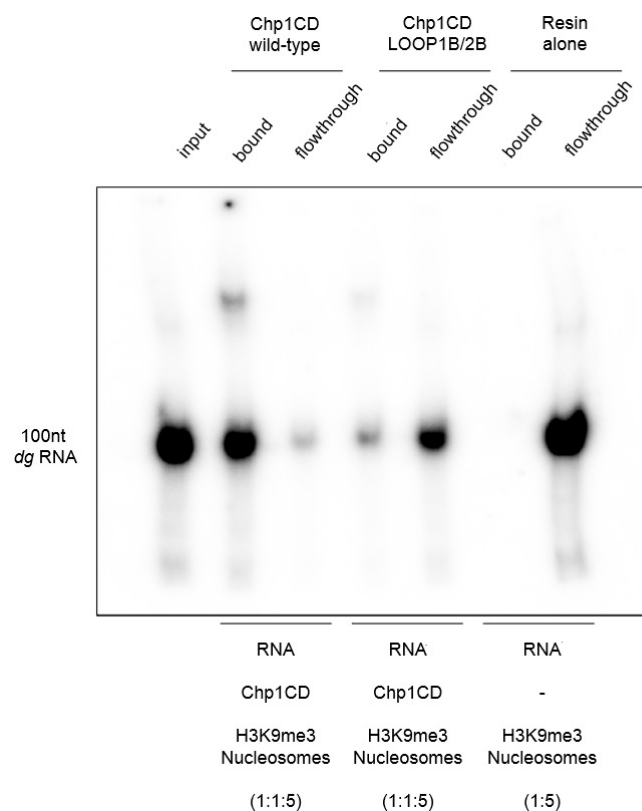


Figure 2.14: RNA pulldowns with Chp1CD (wild-type and LOOP1B/2B mutant) in the presence of H3K9me3 Nucleosomes

³²P radiolabeled 100nt *dg* centromeric RNA pull down experiment showing that the LOOP1B/2B mutants is defective in binding in the presence of H3K9me3 Nucleosomes, compared to the wild-type Chp1CD. The resin control show no non-specific background RNA (bound fraction).

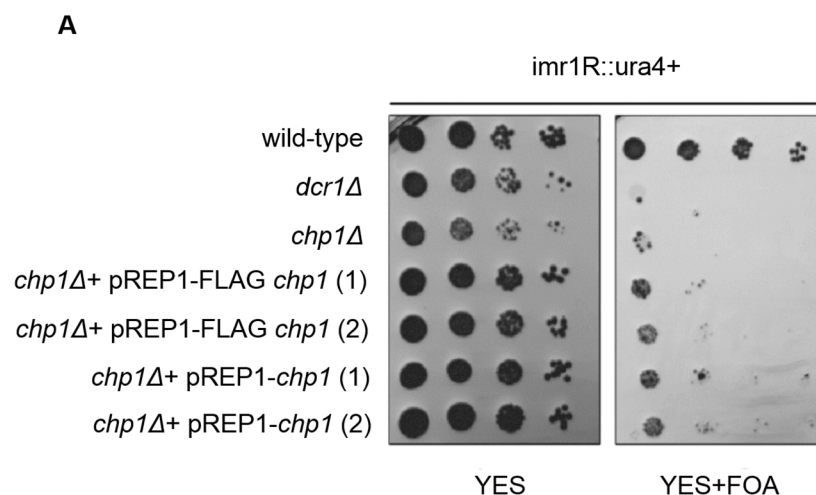
It can be speculated that the impaired binding to the Nucleosome core might interfere with the

Results

ability of Chp1CD LOOP1B/2B to stably interact with the RNA. A potential scenario, in fact, is that both wild-type and LOOP1B/2B mutant Chp1CD pulled-down the same levels of RNA (since the binding of the H3K9me3 tail was still at saturating levels) but that the difference in the overall stability, determined by the Nucleosome core binding, increased the binding interaction strength of the wild-type compared to the LOOP1B/2B mutant. One possibility is that RNA interaction is destabilized in the Chp1CD-Nucleosome core mutants, explaining how the *in vivo* phenotypes obtained are similar to the ones previously described (Ishida et al., 2012) for the DNA/RNA binding mutants (α mut1 and α mut2), with a clear defect in heterochromatin maintenance. This led us to propose a model (Figure 3.1) in which the Chp1CD-Nucleosome core interaction is fundamental for stable anchoring of centromeric RNAs to chromatin, to facilitate their siRNA mediated downstream targeting by the RITS-Argonaute for siRNA amplification, H3K9me2 deposition and heterochromatin formation.

Over-expression of Chp1, Swi6, Chp2 and Clr4 in *S. pombe* cells destabilizes centromeric heterochromatin

The dynamic of interaction between chromodomain proteins in *S. pombe* and heterochromatin is very complex and has to be tightly regulated.



Results

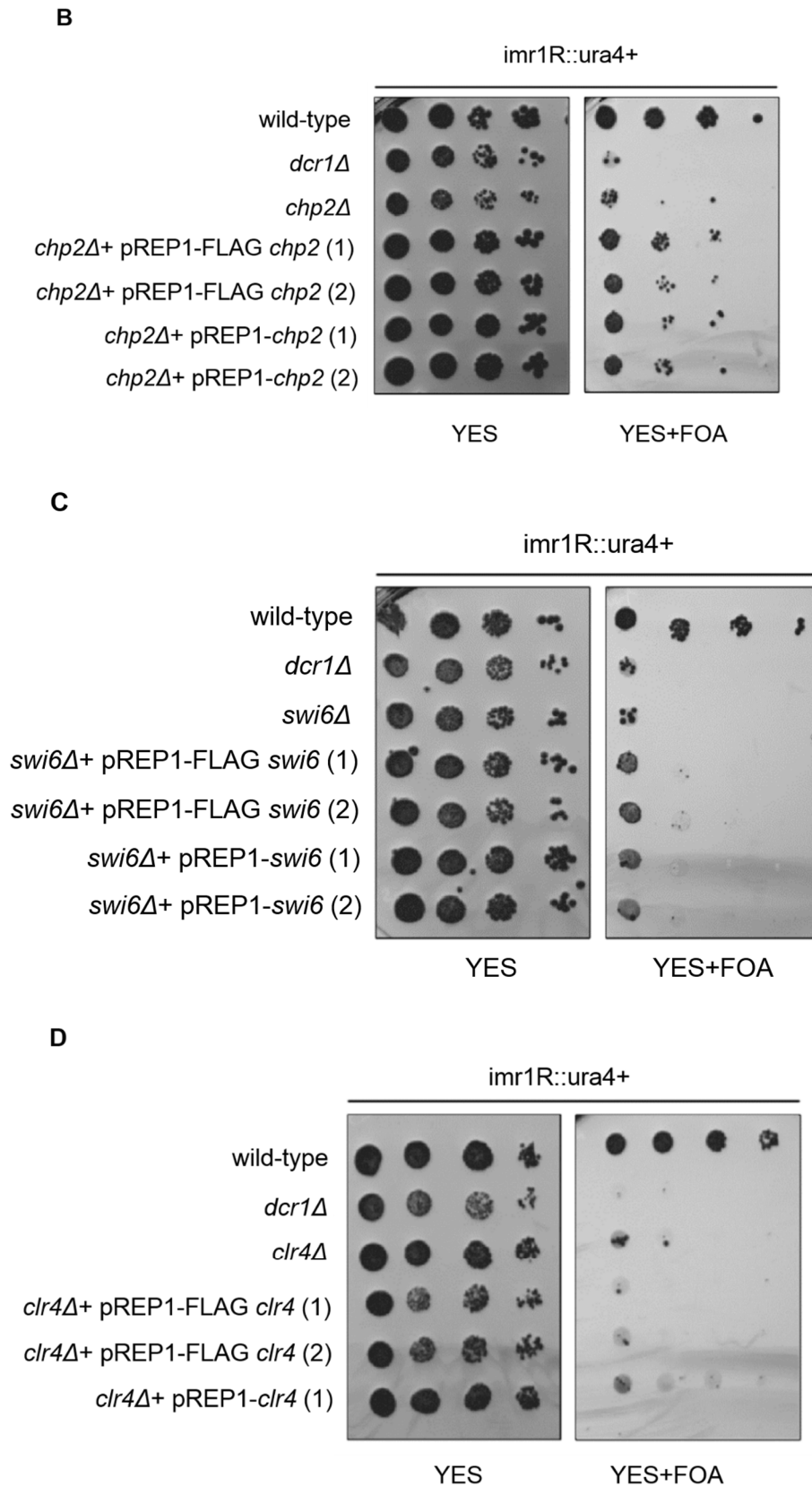


Figure 2.15: over-expression of Chp1, Swi6, Chp2 and Clr4 chromodomain proteins destabilizes centromeric silencing

A) Silencing assay showing that the over-expression of Chp1 under a strong nmt1+ promoter (pREP1 plasmid) cannot rescue the silencing defect of a chp1Δ strain. We have previously showed that a complete plasmid rescue

Results

occurs when Chp1 is expressed under its endogenous promoter (Figure 2.6). 10- fold serial dilutions were spotted on YES (non-selective) and YES+FOA media. Results are shown for two colonies of the Chp1 rescue transformation, for both the FLAG tag and untagged Chp1, indicated as (1) and (2). (B) Silencing assay showing that the over-expression of Chp2 under a strong *nmt1+* promoter (pREP1 plasmid) only partially rescues the silencing defect of a *chp2Δ* strain. 10- fold serial dilutions were spotted on YES (non-selective) and YES+FOA media. Results are shown for two colonies of the Chp2 rescue transformation, for both the FLAG tag and untagged Chp2, indicated as (1) and (2). (C) Silencing assay showing that the over-expression of Swi6 under a strong *nmt1+* promoter (pREP1 plasmid) cannot rescue the silencing defect of a *swi6Δ* strain. 10- fold serial dilutions were spotted on YES (non-selective) and YES+FOA media. Results are shown for two colonies of the Swi6 rescue transformation, for both the FLAG tag and untagged Swi6, indicated as (1) and (2). (D) Silencing assay showing that the over-expression of Clr4 under a strong *nmt1+* promoter (pREP1 plasmid) cannot rescue the silencing defect of a *clr4Δ* strain. 10- fold serial dilutions were spotted on YES (non-selective) and YES+FOA media. Results are shown for two colonies of the Clr4 rescue transformation for the FLAG tag and one colony for the untagged Clr4, indicated as (1) and (2).

It was previously showed that a plasmid carrying the Chp1 protein expressed under its endogenous promoter was able to completely rescue the *chp1Δ* strain silencing defect (Figure 2.6). Therefore, the slight over-expression of Chp1 due to the pREP1 plasmid copy number was not affecting the function of the protein in heterochromatin silencing at centromeres. However, over-expression of Chp1 under an *nmt1+* strong promoter in a *chp1Δ* strain, despite not causing a growth defect, was definitely impairing centromeric heterochromatin formation, and was not rescuing the background *chp1Δ* strain defect (Figure 2.15 A). Similar results were obtained for the HP1 Swi6 (Figure 2.15 C) and for H3K9 methyltransferase Clr4 (Figure 2.15 D). In the case of Chp2, less abundant HP1 isoform of Swi6, over-expression caused a partial rescue of the *chp2Δ* strain centromeric defect (Figure 2.15 B). This finding probably reflects the more specialized function of Chp2 in heterochromatin regulation. Over-expression of Clr4 likely interferes with proper H3K9me mark deposition, thereby altering the equilibrium of HP1 and RITS recruitment across the entire genome. Purification of the over-expressed FLAG tagged version of the proteins, with sufficient yields for structural studies, was successful only for Chp2 (Figure 2.17 A). The *in vivo* purified Chp2 retains its post-translational modifications, which could provide clues on how HP1 chromatin binding is regulated.

To further understand the dynamic of interaction between fission yeast chromodomain proteins and chromatin, additional Nucleosomal complexes were assembled following a similar experimental procedure adopted for Chp1CD, with the aim of solving their structure by using cryo-EM. In the next subchapters, the *in vitro* assembly of H3K9me3 Nucleosomes in a complex with Swi6, Chp2 and Clr4 is discussed.

In vitro reconstitution of the Swi6-H3K9me3 Nucleosome complex

The *S. pombe* fission yeast HP1 homolog Swi6 was expressed and purified from *E.coli* BL21(DE3)pLysS bacterial strain from a pETDuet (Novagen) vector as an His₆-Swi6-FLAG

Results

double tag construct. The purification protocol followed was the same as in the case of the His₆-SUMO-Chp1CD: we purified His₆-Swi6-FLAG through Ni-NTA affinity purification, as described for Chp1CD. As showed in Figure 2.16 A, the His₆-Swi6-FLAG runs between 50 and 60 kDa, higher than the calculated molecular weight of 37.29 kDa. After Ni-NTA purification, His₆-Swi6-FLAG was bound to the FLAG M2 affinity gel in Nucleosome binding buffer. With the FLAG resin binding step, we were able to further purify Swi6 from contaminants and degradation products. Once Swi6 was bound to the FLAG resin, we assembled the Swi6-H3K9me3 Nucleosome complex by adding H3K9me3 Nucleosomes, following the experimental scheme shown in Figure 2.16 B. We then eluted the complex from the resin by adding FLAG peptide. The addition of FLAG peptide was able to competitively elute the Swi6-H3K9me3 Nucleosome complex (Figure 2.16 C - fourth lane from the left). If no FLAG peptide was added, the Swi6 Nucleosome complex remained stably anchored to the resin (Figure 2.16 C - fifth lane from the left). Moreover, no H3K9me3 Nucleosome was detected when no Swi6 was bound to the FLAG resin, indicating that the complex formation observed was specific (Figure 2.16 C - second lane from the right). We then performed Negative stain EM on the eluted Swi6-H3K9me3 Nucleosome complex to verify its quality and homogeneity (Figure 2.16 D), as previously done for the Chp1CD-H3K9me3 Nucleosome complex (the procedure is described in Materials and Methods). Particles were homogeneous and concentrated enough to cover the majority of the carbon coated EM grid holes. Differently from what observed for Chp1CD-H3K9me3 Nucleosome, the Swi6-H3K9me3 Nucleosome complexes seemed to have outwards protrusions and connections with each other, which might reflect the way Swi6 binds to Nucleosomes and oligomerizes. On the contrary, negative stain EM can give rise to artifacts, derived from the addition of the heavy metal Uranyl acetate stain, causing protein flattening and non-specific aggregation. It would be therefore premature to make any interpretations. Further structural studies using cryo-EM and new generation electron detectors will be needed to achieve high resolution and to understand how Swi6 interacts with H3K9me3 Nucleosomes under physiological conditions. Nonetheless, the preparation of the Swi6-H3K9me3 Nucleosome complex for EM studies represents an important step toward the structural characterization of the HP1- Nucleosome interaction.

Results

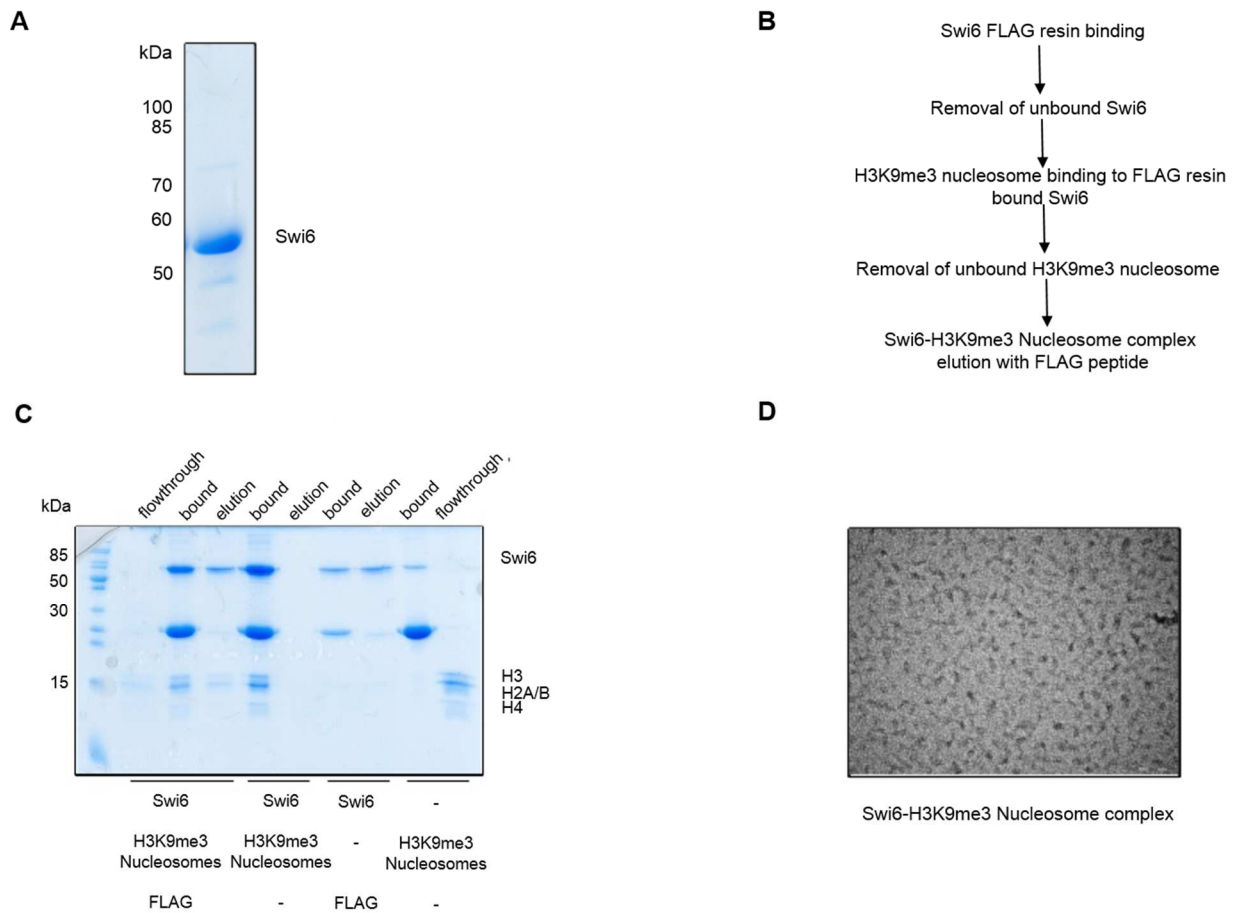


Figure 2.16: *in vitro* assembly of the Swi6-H3K9me3 Nucleosome complex

A) Coomassie stained gel showing the purified His-Swi6-FLAG protein from *E. coli* BL21(DE3)pLys expression strain. The protein runs at around 55 kDa. *(B)* Experimental scheme for the Swi6-H3K9me Nucleosome complex assembly. *(C)* Coomassie stained 15% SDS-PAGE gel showing the assembly of Swi6-H3K9me3 Nucleosome complex and its elution from the FLAG resin (Sigma) upon FLAG peptide (Sigma) addition. *(D)* Negative stain EM micrograph showing the Swi6-H3K9me3 Nucleosome complex.

In vitro reconstitution of the Chp2-H3K9me3 Nucleosome complex

The *S. pombe* Chp2 was purified from fission yeast *chp2Δ* cells transformed with a pREP1 plasmid overexpressing an N-terminal FLAG tagged Chp2 under the *nmt1* + promoter (Figure 2.15 B). As discussed in the Introduction, Chp2 is the HP1 Swi6 isoform, with similar heterochromatin localization and almost overlapping functions. The advantage of the *in vivo* purification approach used is that Chp2 maintains its post-translational modifications (PTMs) and this could give more insights on how the binding to chromatin is regulated. However, over-expression might change not only the cellular localization of Chp2 but also affect the type of PTMs deposited on its residues. The structural characterization of the *in vivo* purified Chp2 in complex with Nucleosomes has therefore to be coupled with proper identification of Chp2

PTMs, in the wild-type and overexpression strain, to better understand the regulation of its binding. FLAG-Chp2 was immuno-precipitated by using the FLAG M2 affinity gel. Like in the case of Swi6, Chp2 runs at roughly 65 kDa, a higher molecular weight than the one calculated from the primary sequence: 42.95 kDa (Figure 2.17 A). The resin was washed to remove the unbound fraction and the Chp2-H3K9me3 Nucleosome complex was assembled by incubating the resin bound FLAG-Chp2 with H3K9me3 Nucleosomes. The experimental scheme is shown in Figure 2.17 B. After washing the unbound H3K9me3 Nucleosomes from the FLAG resin, the Chp2-H3K9me3 Nucleosome complex was eluted by addition of the Prescission protease. The FLAG-Chp2 construct has in fact a Prescission protease digestion site between the FLAG tag sequence and the N-terminal start of the protein. Addition of the Prescission protease, removed the tag anchoring Chp2 to the resin, thereby eluting the Chp2-H3K9me3 Nucleosome complex (Figure 2.17 C- fourth lane from the left). The advantage of the Prescission lies in its high specificity and in the low optimal temperature of 4 °C, which allowed to complete the elution on ice, without destabilizing the Chp2-H3K9me3 Nucleosome complex. Moreover, specificity was required since many impurities were still present on the resin after FLAG-Chp2 immuno-precipitation, which would have been carried over on the EM grid if we would have performed a general FLAG peptide elution. Controls on the complex assembly, showed that elution happened only when adding the Prescission protease (Figure 2.17 C- fifth lane from the left) and that no H3K9me3 Nucleosome was non-specifically sticking to the FLAG resin (Figure 2.17 C- second lane from the right). Negative stain EM of the eluted complex showed the presence of particles with similar shape and size, even though less concentrated if compared to what previously achieved with the Swi6-H3K9me3Nucleosome complex (Figure 2.17 D). A similar Swi6 dual tag approach should be used in the future also for the Chp2-H3K9me3 Nucleosome complex, moving to a bacterial expression system, which generally gave better protein yields for Chp1CD and Swi6 proteins. Biochemical and structural characterization of the *in vivo* Chp2-H3K9me3 Nucleosome complex might be continued in parallel, since maintaining post-translational modifications on the protein could give further insights on the regulation of HP1 binding to chromatin. However, the elution strategy has to be optimized to give a more concentrated EM sample.

Results

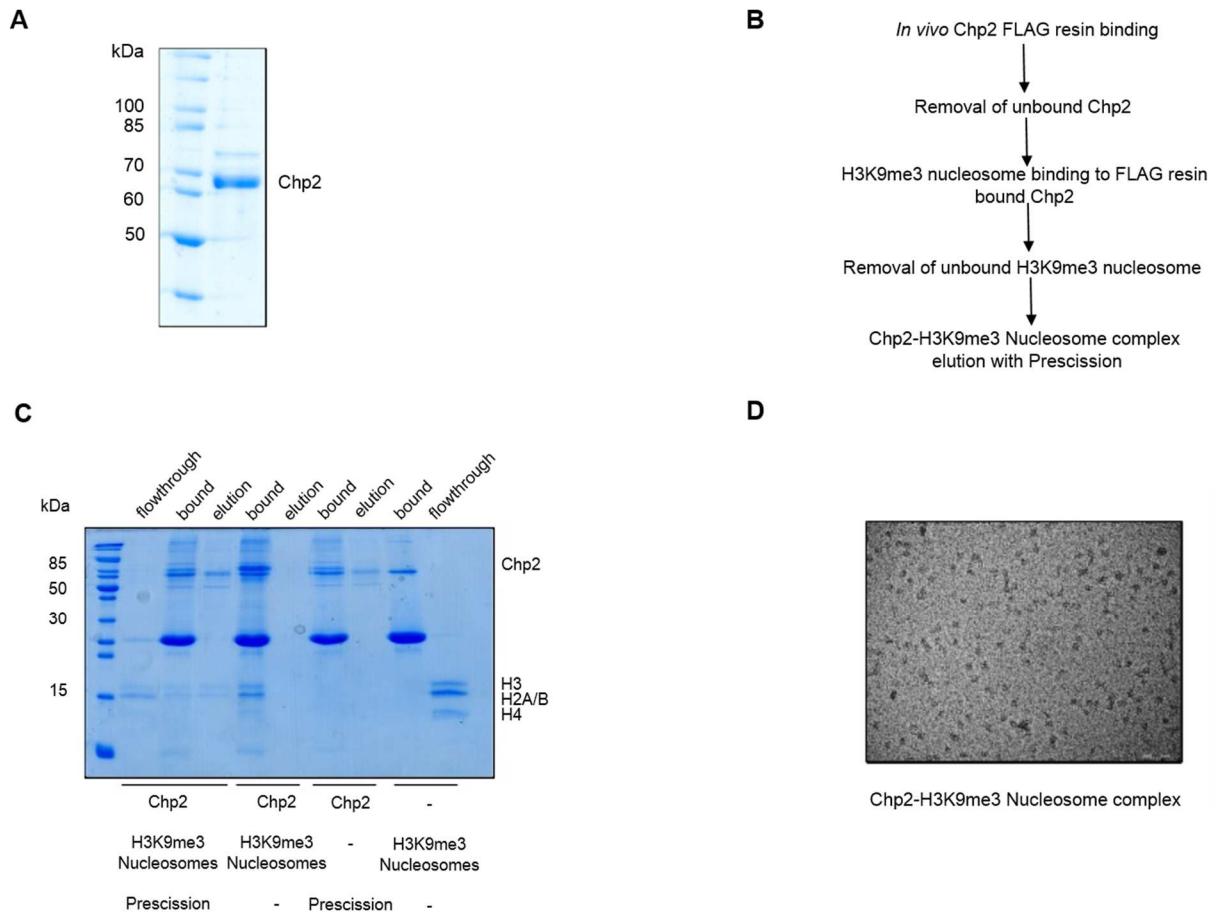


Figure 2.17: *in vitro* assembly of the Chp2-H3K9me3 Nucleosome complex

A) Coomassie stained gel showing the purified FLAG-Chp2 protein from a wild-type *S.pombe* strain. The strain was transformed with a pREP1 plasmid over-expressing FLAG-Chp2 under the *nmt1+* promoter. The protein runs at around 65 kDa. *(B)* Experimental scheme for the Chp2-H3K9me Nucleosome complex assembly. *(C)* Coomassie stained 15% SDS-PAGE gel showing the assembly of Chp2-H3K9me3 Nucleosome complex and its elution from the FLAG resin (Sigma) upon addition of the Prescission protease. A Prescission cut site is present between the FLAG tag and the Chp2 protein N-terminal end. *(D)* Negative stain EM micrograph showing the Chp2-H3K9me3 Nucleosome complex.

In vitro reconstitution of the Clr4-H3K9me3 Nucleosome complex

The *S. pombe* H3K9 methyltransferase Clr4 was purified from the *E. coli* BL21(DE3)pLysS expression strain as a double tag construct His₆-Clr4-FLAG (we thank Bassem Al-Sady for the construct, published in (Al-Sady et al., 2013)). The purified protein run at around 80 kDa, higher than the 55.92 kDa calculated based on its primary sequence (Figure 2.18 A). As described previously for Swi6, the assembly of the Clr4-H3K9me3 Nucleosome complex followed the procedure described in Figure 2.18 B. After the initial Ni-NTA purification, Clr4 was rebound to the FLAG M2 affinity resin. Then H3K9me3 Nucleosomes were added to assemble the Clr4-H3K9me3 Nucleosome complex. The binding occurred in the same buffer conditions described

for Chp1CD, Swi6 and Chp2. After washing the unbound H3K9me3 Nucleosomes, the Clr4-H3K9me3 Nucleosome complex was eluted by addition of the FLAG peptide (Figure 2.18 C - fourth lane from the left). Complex elution was dependent on addition of the FLAG peptide, and no Nucleosome was spontaneously dissociating from the resin (Figure 2.18 C - fifth lane from the left). As in the case of Swi6, Clr4-H3K9me3 Nucleosome complex assembly was specific and no Nucleosome was observed sticking on the FLAG resin in the absence of bound Clr4 (Figure 2.18 C - second lane from the right). The eluted complex was later spotted on a carbon coated EM grid for Negative staining. We observed homogenous and well distributed Clr4-H3K9me3 Nucleosome particles, with a round shape and similar size (Figure 2.18 D). As for the other Nucleosome complexes purified and assembled, cryo-EM data on the assembled Clr4-H3K9me3 Nucleosome complex will be collected in order to better understand how the Clr4 H3K9 methyltransferase interacts with the Nucleosome and exerts its own enzymatic function. This could potentially give precious informations on how the H3K9me mark is deposited and how the chromodomain of Clr4, which shares striking primary sequence similarity with Chp1CD, binds to the H3 histone tail. Clr4 is of particular interest since, among the four fission yeast heterochromatin proteins with a chromodomain, it is the only one coupling the “chromatin reader” and the “chromatin writer” functions. A high resolution structure of this complex will be therefore beneficial to the chromatin field, given also the conservation of the Su(var)3-9 family of H3K9 methyltransferases and their fundamental role in heterochromatin formation.

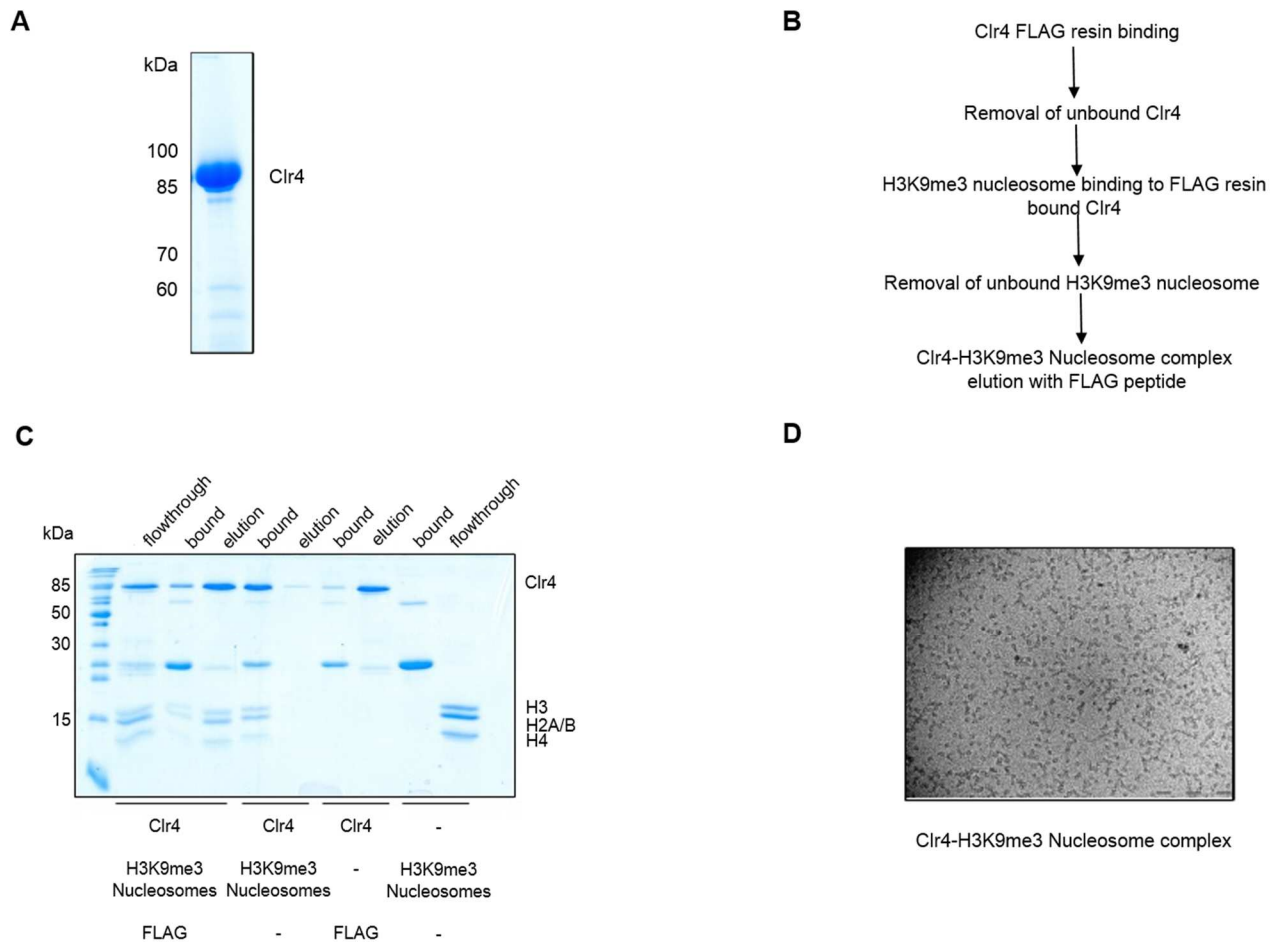


Figure 2.18: *in vitro* assembly of the Ctr4-H3K9me3 Nucleosome complex

A) Coomassie stained gel showing the purified His-Ctr4-FLAG protein from *E.coli* BL21(DE3)pLys expression strain. The protein runs at around 80 kDa. *(B)* Experimental scheme for the Ctr4-H3K9me Nucleosome complex assembly. *(C)* Coomassie stained 15% SDS-PAGE gel showing the assembly of Ctr4-H3K9me3 Nucleosome complex and its elution from the FLAG resin (Sigma) upon FLAG peptide (Sigma) addition. *(D)* Negative stain EM micrograph showing the Ctr4-H3K9me3 Nucleosome complex.

Cross-linking of the Swi6-H3K9me3 Nucleosome complex

In order to better stabilize the Swi6-H3K9me3 Nucleosome interaction for structural studies and Mass Spectroscopy, cross-linking was performed on the assembled complex by using bis(sulfosuccinimidyl)suberate or BS³. BS³ is a homobifunctional amine-amine cross-linker that reacts with two primary amine residues, such as lysines, covalently linking them together with a spacer distance of 11.4 Å. In order to assess the optimal BS³ concentration for cross-linking, avoiding over reaction, titration experiments were set up, using Swi6 and Nucleosomes alone and the Swi6-H3K9me3 Nucleosome complex (Figure 2.19). Together with Swi6, the SwiL315D protein (Canzio et al., 2011), carrying a mutation in the Chromo shadow domain preventing dimerization in solution, was analyzed.

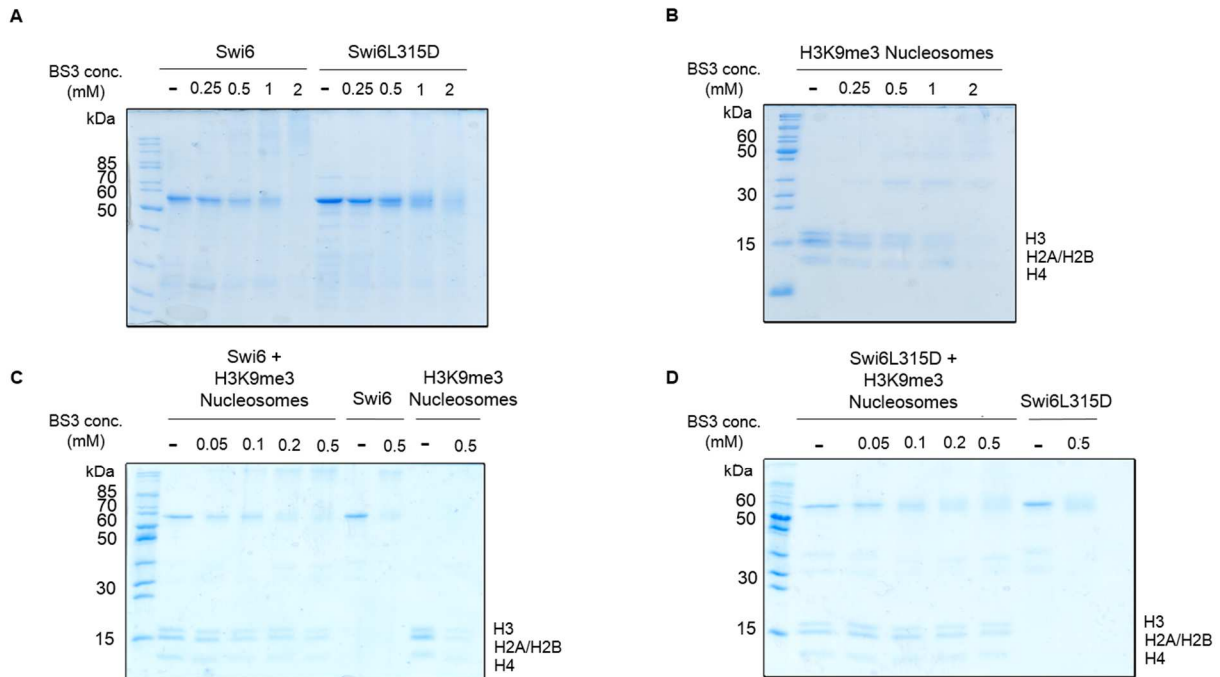


Figure 2.19: BS³ cross-linking titration experiments for the Swi6-H3K9me3 Nucleosome complex

A) Coomassie stained 12% acrylamide gel of a BS³ titration experiment on the Swi6 protein and the Swi6L315D dimerization mutant. As can be noticed, at higher concentration of cross-linker (1 and 2 mM) Swi6 but not Swi6L315D, shifts upwards, indicating the formation of oligomers. (B) Coomassie stained 15% acrylamide gel of a BS³ titration experiment on the H3K9me3 Nucleosome. Histone bands are visible up to 1 mM BS³ concentration. (C) Coomassie stained 15% acrylamide gel of a BS³ titration experiment on the Swi6-H3K9me3 Nucleosome complex, assembled in solution. (D) Coomassie stained 15% acrylamide gel of a BS³ titration experiment on the Swi6L315D-H3K9me3 Nucleosome complex, assembled in solution.

The analysis on the Swi6L315D mutant is fundamental to understand the contribution of Swi6 oligomerization for Nucleosome interaction. The cross-linking reaction was carried out at 30°C for 30 minutes and quenched by the addition of 50 mM Ammonium Carbonate for 10 minutes, for a total protein concentration of 0.5 mg/ml.

As shown in Figure 2.19 A, a concentration of BS³ of 0.5 mM or higher caused a shift upwards of the Swi6 protein in SDS-PAGE, indicating a covalent bond formation between Swi6 dimers. As expected, the Swi6L315D mutant, unable to dimerize, did not show the same shifting (Figure 2.19 A). H3K9me3 Nucleosomes started showing higher molecular weight histone bands starting from a BS³ concentration of 0.5 mM (Figure 2.19 B). In conclusion the minimal BS³ concentration for cross-linking the Swi6-H3K9me3 Nucleosome complex, without over reaction, was around 0.5 mM. We assembled the Swi6-H3K9me3 Nucleosome and the Swi6L315D-H3K9me3 Nucleosome complexes in solution and performed low concentration BS³ titration experiments, starting from 0.05 mM to 0.5 mM BS³ (Figure 2.19 C and D). The

Results

appearance of cross-linked molecular species occurred at a concentration of BS³ comprised between 0.1 mM and 0.5 mM. For cross-linking Mass Spectrometry, the minimal concentration of cross-linker able to efficiently cross-link H3K9me3 Nucleosomes and the Swi6-H3K9me3 Nucleosome complex was used, namely 0.5 mM. The Swi6 protein alone, but not Nucleosomes, shifted at lower BS³ concentrations.

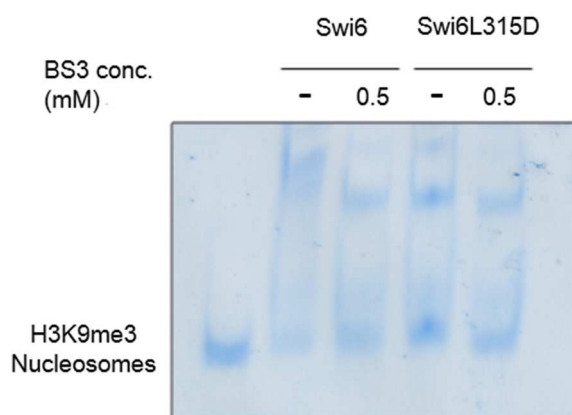


Figure 2.20: cross-linking of the Swi6-H3K9me3 Nucleosome complex

Native-TBE acrylamide gel showing the cross-linked Swi6-H3K9me3 Nucleosome and the Swi6L315D-H3K9me3 Nucleosome complex with and without BS³ cross-linking (0.5 mM). Note the upper shifted Nucleosome band corresponding to the assembled Swi6-Nucleosome complex. 0.5 mM BS³ cross-linking does not change the migration of the Swi6-Nucleosome complex on Native PAGE.

As shown in Figure 2.20, cross-linking with a concentration of 0.5 mM BS³ did not change the migration properties of the Swi6-H3K9me3 Nucleosome complex in Native PAGE.

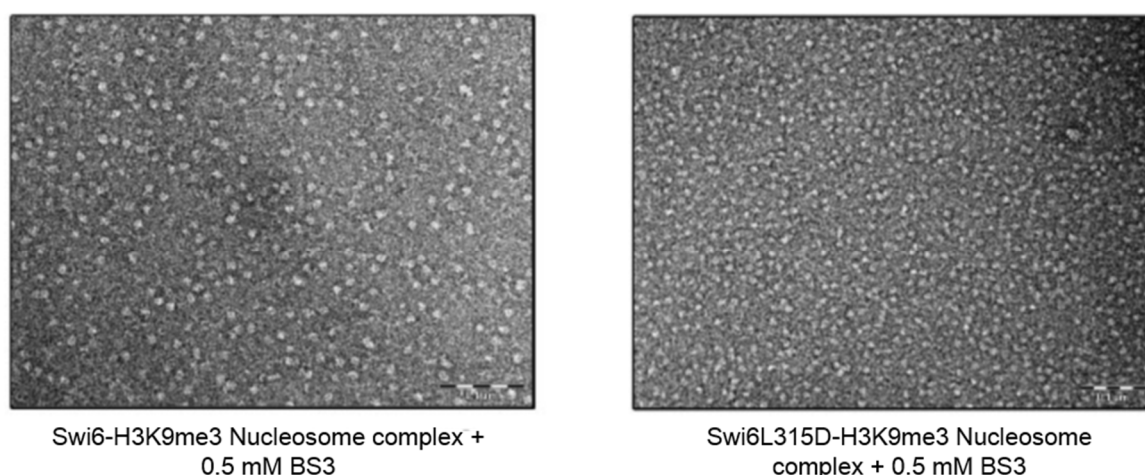


Figure 2.21: Negative stain EM of the cross-linked Swi6-H3K9me3 Nucleosome and the Swi6L315D-H3K9me3 Nucleosome complexes

In fact, in EMSA with or without 0.5 mM BS³, there was a clear upward shifting Nucleosomal band, corresponding to the assembled Swi6-H3K9me3 Nucleosome complex. The cross-linked Swi6 and Swi6L315D Nucleosomal complexes were also subjected to Negative stain EM

Results

(Figure 2.21), to assess the quality of the particles. On the EM grids, cross-linked complexes looked similar in size and in shape, with an optimal concentration to cover homogenously the carbon surface. The sample seemed, therefore, to be optimal to proceed further with the cryo-EM structural analysis.

Cross-linking Mass Spectrometry data will allow us to pin-point the residues in both Swi6 and the Swi6L315D dimerization mutant directly interacting with the histone H3K9me3 tail and the Nucleosome core. This powerful method, coupled with cryo-EM, will give us an idea on how Swi6 structurally interacts with Nucleosomes to form heterochromatin.

Reconstitution of H3K9me3 di-Nucleosomes

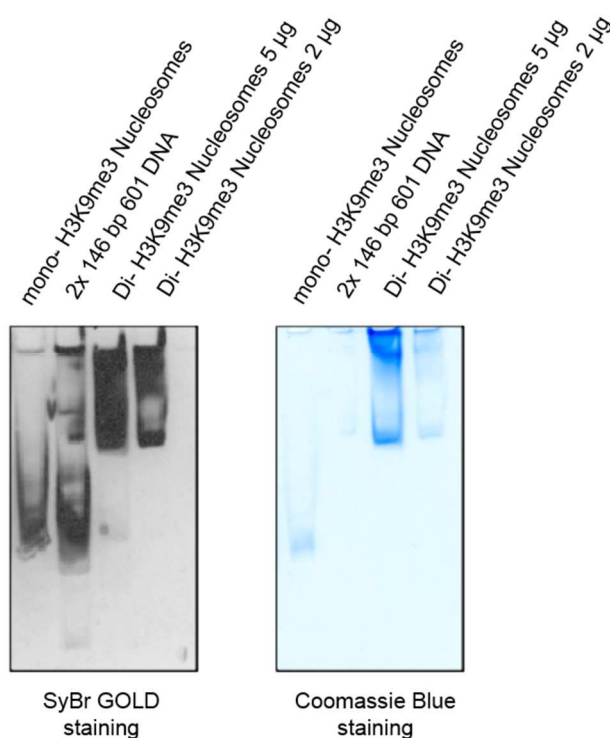


Figure 2.22: *in vitro* reconstitution of H3K9me3 di-Nucleosomes

H3K9me di-Nucleosome reconstitution: nucleosomes were run on a 5% TBE-Native acrylamide gel and stained with SyBr GOLD and Coomassie Blue. Mono H3K9me3 Nucleosomes (NCP) were loaded as a control. Two different quantities (5 µg and 2 µg) of di-H3K9me3 Nucleosomes were run. Note the shifting of the 2x146bp 601 DNA in the di-Nucleosome samples and the double staining with SyBr GOLD and Coomassie Blue.

To further expand the analysis on chromodomain proteins chromatin interaction, H3K9me3 di-Nucleosomes were reconstituted for cryo-EM studies. As substrate for reconstitution we used

Results

the 2 x 146 bp 601 DNA sequence, containing two copies of the 601 146 bp Widom sequence previously used to assemble single Nucleosome core particles (Lowary and Widom, 1998). The di-H3K9me3 Nucleosomes were reconstituted following the same procedure described previously for the mono H3K9me3 Nucleosomes. Histone Octamer was mixed with the 2x 146 bp 601 DNA sequence in a DNA : Octamer ratio of 1.1:1, in a Step gradient dialysis lowering the salt concentration. As shown in Figure 2.22, the reconstituted di-H3K9me3 Nucleosomes run higher compared to the 2 x 146 bp 601 DNA and the mono- H3K9me3 Nucleosomes alone and were stained with both SyBr GOLD and Coomassie Blue. To understand whether assembled di-H3K9me3 Nucleosomes were functional for Swi6 binding an EMSA was performed with Swi6 - di-H3K9me3 Nucleosomes in a 5 : 1 ratio, in both cross-linked (0.5 mM BS³) and native conditions (Figure 2.23).

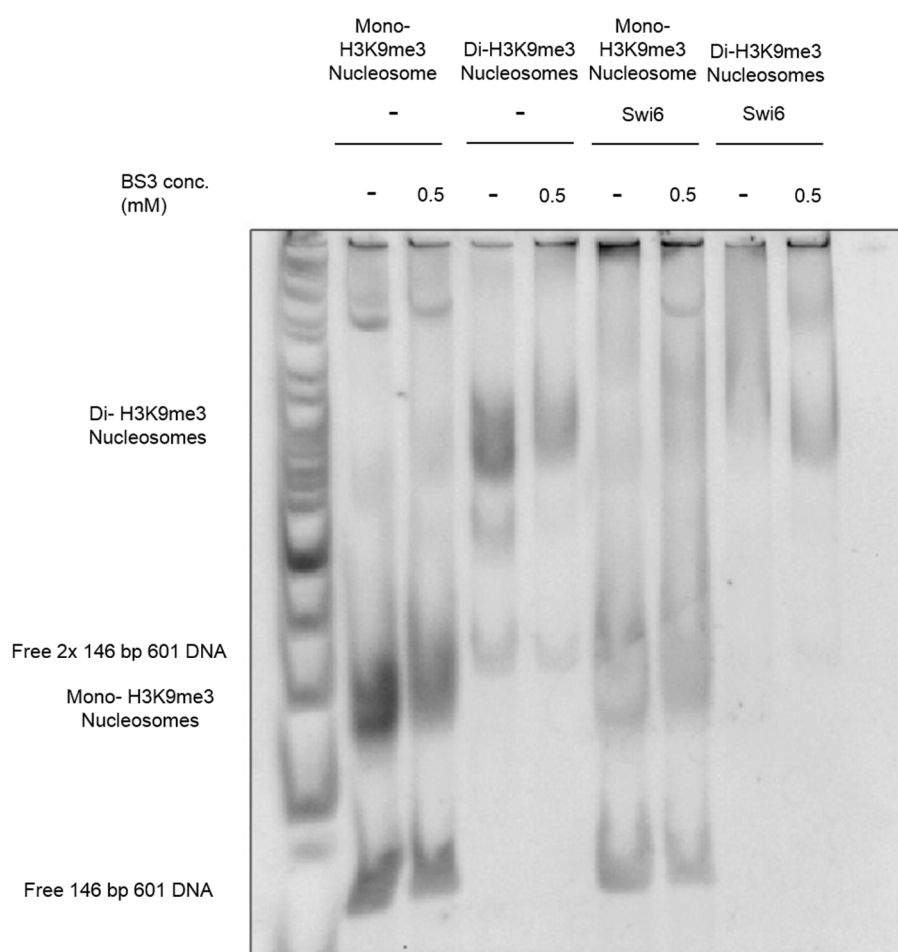


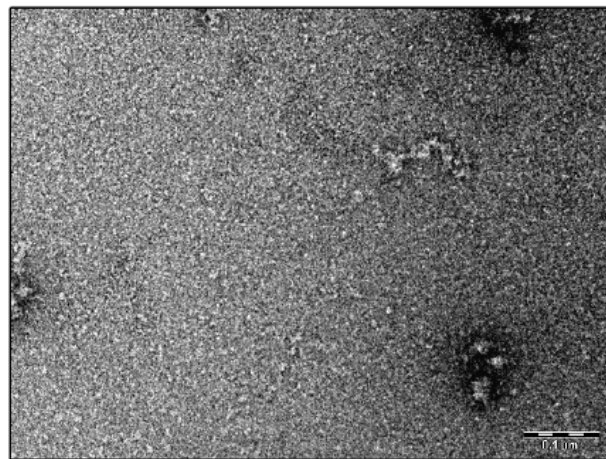
Figure 2.23: Swi6-H3K9me3 di-Nucleosome complex assembly

Native -TBE acrylamide gel showing the cross-linked and native Swi6-H3K9me3 mono-Nucleosome and the Swi6-H3K9me3 di- Nucleosome complex with and without BS³ cross-linking (0.5 mM).

The addition of Swi6 to H3K9me3 di-Nucleosomes caused a shifting in the Nucleosomal band,

Results

indicating Swi6 binding. Addition of 0.5 mM BS³ stabilized the Swi6 - H3K9me3 di-Nucleosome complex, with the appearance of a more defined shifted Nucleosomal band (Figure 2.23 – compare the last two lanes). Since it was possible to assess the binding of Swi6 to H3K9me3 di-Nucleosomes, we next wanted to assess the quality of the assembled di-Nucleosome particles by Negative stain EM.



Di-H3K9me3 Nucleosomes

Figure 2.24: Negative stain EM micrograph of reconstituted H3K9me3 di-Nucleosomes

H3K9me3 di-Nucleosomes showed the presence of many protein aggregates compared to mono- H3K9me3 Nucleosomes in the same buffer, with many debris and irregular shaped particles. The addition of 1 mM EDTA to the Nucleosome binding buffer, did not resolve the H3K9me3 di-Nucleosome aggregation on the Negative stain EM grid (Figure 2.24). It is fundamental in order to use di-Nucleosomes for cryo-EM studies, to understand how to prevent aggregation, which might be derived from sensitivity to the Uranyl Acetate low pH, from carbon interaction that causes di-Nucleosome disassembly or general biochemical instability. Future optimization of di-Nucleosome chromodomain protein complexes will shed more light on how heterochromatin is assembled and regulated in complex chromatin substrates.

DISCUSSION

The organization of the chromatin into euchromatin and heterochromatin is fundamental to regulate a wide variety of nuclear processes, such as transcription, DNA replication and chromosome segregation, DNA repair, genome organization in the nucleus and its expression. In *S. pombe* as well as in many other Eukaryotes, the recognition of H3K9 methylation from chromodomain proteins is essential to establish and maintain constitutive heterochromatin, and to recruit protein complexes involved in the regulation of various nuclear processes. However, there is still no sufficient insights on how heterochromatin chromodomain proteins interact with their natural binding partners: the H3K9me_{2/3} Nucleosomes. There are still many open questions on how the different chromodomain proteins involved in heterochromatin formation recognize the H3K9 methylation mark and cross-talk with each other to coordinate their functions at the same genomic locus.

In the present work of thesis the attention was mainly focused on the chromodomain of Chp1, a protein involved in the recruitment of Argonaute to chromatin, to trigger the RNA interference mediated heterochromatin formation in the fission yeast *S. pombe*. Among the four chromodomain proteins involved in heterochromatin formation in fission yeast (Chp1, Chp2, Swi6 and Clr4), the chromodomain of Chp1 is the one with the highest affinity for the H3K9 methylation mark. This was already shown previously by various *in vitro* H3K9me peptide binding experiments (pull downs, fluorescence anisotropy and isothermal calorimetry-ITC), which, even despite showing differences in the affinity constants (K_d) measured, they consistently underlined that Chp1CD binds the H3K9me₃ tail with the highest affinity (Al-Sady et al., 2013; Ishida et al., 2012; Schalch et al., 2009). Ishida and colleagues (Ishida et al., 2012), measured by ITC that the affinity of the Chp1CD for the H3K9me₃ peptide was around 37.5 nM, whereas Swi6CD affinity was roughly 160 nM. Schalch and colleagues (Schalch et al., 2009) measured an H3K9me₃ peptide affinity by fluorescence anisotropy of roughly 190 nM for Chp1CD, 600 nM for Clr4CD and 3.34 μM for the Swi6CD. However, isothermal calorimetry experiments gave slightly less difference in the binding affinity between Chp1CD and Swi6, with Chp1CD binding the H3K9me₃ peptide with a K_d = 0.3 μM and Swi6 with a K_d = 1.69 (Schalch et al., 2009). For Zhang and colleagues, the measured affinity by fluorescence polarization of the Clr4CD for the H3K9me₃ peptide was 6 μM (Zhang et al., 2008). In a recent study (Al-Sady et al., 2013), the Clr4 protein with the Clr4CD swapped with a F61A Chp1CD (Schalch et al., 2009), was still able to recognize the H3K9me₃ mark with the same affinity of the Clr4CD, but had tighter binding to H3K9me₂. This was done to prevent

the swapped chimeric F61AChp1CD-Clr4 protein to discriminate between H3K9me3 and H3K9me2 marks. It was highlighted however that the wild-type Chp1CD had higher affinity compared to the Clr4CD for the H3K9me mark, independently of the methylation state, as previously measured (Schalch et al., 2009). In this study, the H3K9me3 peptide affinities measured by Microscale Thermophoresis (MST) of Chp1CD chromodomain wild-type (K_d = 345 nM) and the LOOP1B (K_d = 260 nM), LOOP2B (K_d = 292 nM) and LOOP1B/2B (K_d = 365 nM) mutants, are all in a comparable range to what previously shown in ITC measurements for the Chp1CD construct (Schalch et al., 2009).

However, no experiments have been done so far to compare the binding affinity of all the chromodomains mentioned with H3K9me3 Nucleosomes *in vitro*. In this study, MST measurements of the wild-type Chp1CD affinity in complex with the H3K9me3 Nucleosome gave a dissociation constant of K_d = 180 nM, a value lower than the one measured with the H3K9me3 peptide (K_d = 345 nM). This can be due to the presence of additional Nucleosomal surface that can increase the affinity of binding, even though these values are difficult to compare since the binding buffer used in the Nucleosome binding assay had to be supplemented (+10% glycerol compared to the buffer used in H3K9me3 peptide assays, see Materials and Methods)) to prevent sample aggregation on the thermophoresis capillaries.

What was already shown, on the other hand, is that a heterochromatin defective phenotype is obtained when swapping chromodomains between Chp1, Swi6, Chp2 and Clr4.

Swapping of Swi6CD with its HP1 isoform Chp2CD caused a major defect in silencing at the MAT locus (Sadaie et al., 2008). On the other hand, swapping of Chp2CD with Swi6CD had a less impactful heterochromatin defect. The swapping of Chp1CD with Swi6CD resulted in defective heterochromatin establishment at centromeres, with no effects on heterochromatin maintenance (Schalch et al., 2009). The swapping previously described (Al-Sady et al., 2013) of the Clr4CD with a less H3K9me2/3 specific F61AChp1CD, caused impaired chromatin spreading at the MAT type locus. Moreover, the swapping of the Swi6CD with either Chp1CD or Clr4CD caused a slight defect of silencing at centromeres (Ishida et al., 2012). No defect was observed when swapping Swi6CD with Chp2CD at centromeres, in contrast to what previously seen at the MAT type locus (Sadaie et al., 2008). This might hint on the fact that Swi6 and Chp2 have equivalent functions at centromeres but not at the MAT locus, despite chromatin fractionation experiments (Sadaie et al., 2008) showed that Swi6 is mainly localized in the soluble fraction (Nucleoplasm), whereas Chp2 is more in the chromatin pellet fraction (Chromatin bound), and that the localization of the two proteins is largely determined by their Chromo-Shadow domains (CSD). The cited references provided useful insights in

understanding how affinity regulates the different roles of chromodomain proteins at the same loci. Particularly interesting is the model proposed by Al-Sadi and colleagues in which the stronger affinity of Clr4 for H3K9me3 mark compared to the H3K9me2 mark allows Clr4 to outcompete HP1 proteins for the H3K9me3 mark. On the other hand, HP1 proteins have a stronger affinity for the H3K9me2 mark than Clr4, which is the most abundant H3K9 methylation variant present *in vivo* in fission yeast. The model presented (Al-Sady et al., 2013) explains how the careful balance between H3K9me deposition by Clr4 and the differential affinities for the H3K9me variants contribute to regulate the function of the chromodomain proteins in heterochromatin formation.

The different affinities of the four *S. pombe* heterochromatin chromodomain proteins for H3K9 methylation are definitely dependent on the CDs primary amino acid sequences, rather than their fold (which is the conserved canonical chromodomain fold, with a three stranded anti-parallel β -sheet and C-terminal α -helix (Horita et al., 2001; Ishida et al., 2012; Schalch et al., 2009). The mentioned CD swapping experiments indicated that each chromodomain has peculiar characteristics that determine the specific functions of each heterochromatin protein and that they are therefore not interchangeable, despite recognizing the same target, the H3K9me histone tail.

A preliminary cryo-EM structure of the Chp1CD-Nucleosome complex obtained in our laboratory (data not shown) revealed that Chp1CD is actually interacting not only with the H3K9me tail but also with the Nucleosome core, through protein-protein direct interaction mediated by the two loops of its three stranded antiparallel β -sheet. The mutagenesis study highlighted that destabilization of either loops (LOOP1, LOOP2 mutants or LOOP1/2 mutants) causes partial or total loss of the Nucleosome core interaction with consequent destabilization of the Chp1CD-H3K9me3 Nucleosome complex, resulting in an affinity drop of roughly 5 fold compared to the wild-type Chp1CD (see Figure 2.3 E, LOOP1B/2B mutant). As published previously (Schalch et al., 2009), genomically integrated Chp1CD mutants with similar affinity defects for H3K9me3 peptide binding did not have a visible phenotype in heterochromatin maintenance, but they were defective in heterochromatin establishment. In establishment assays, the sole H3K9me methyltransferase Clr4 is re-introduced in a Clr4 deletion strain. The mutants of the protein of interest are generated in the Clr4 deletion strain and, if heterochromatin cannot be re-established after Clr4 re-introduction, the mutant protein causes a defective heterochromatin establishment. With the re-introduction of Clr4 and background levels of H3K9me mark, the high affinity of Chp1CD for the H3K9me mark is required for the

recruitment of the RITS-Argonaute complex to give the initial signal for heterochromatin formation and generate high levels of siRNAs (Halic and Moazed, 2010). In maintenance, the presence of other back-up mechanisms, including the Ctr4 read-write maintenance function (Audergon et al., 2015; Ragunathan et al., 2014), makes the affinity of Chp1CD for the H3K9me mark less critical to feed-in the heterochromatin signaling. In our case, the general loss of affinity *in vitro* for the H3K9me3 Nucleosome in our LOOP1/2 mutants reflected immediately in a defective silencing in heterochromatin maintenance, both in Chp1 deletion strains transformed with plasmids carrying wild-type and mutant LOOP1/2 Chp1CD proteins (heterochromatin rescue assays) and in genomically integrated LOOP1/2 Chp1CD mutants. Consequently, a defect in heterochromatin maintenance reflects a defect in heterochromatin formation on a general level, since cells cannot keep optimal H3K9me levels and constitutive heterochromatin silencing throughout the generations. Comparison between the heterochromatin defects obtained in our genomically integrated versus the plasmid rescued strains highlights that the slight over-expression of the Chp1 protein in the plasmid rescue strains *in vivo* is not responsible for the phenotypes observed. Moreover, the mutant Chp1CD proteins have been showed to be stably expressed (Figure 2.10), and the LOOP1B/2B Chp1CD mutant, showing a strong heterochromatin defect *in vivo*, is functional (therefore LOOP1B/2B has to be properly folded) in binding the H3K9me3 histone peptide with an affinity comparable to the wild-type Chp1CD. LOOP1B/2B Chp1CD mutant showed the importance of the Chp1CD-Nucleosome core interaction in heterochromatin formation. Despite having only a 5-fold drop in affinity for the H3K9me3 Nucleosome (considered not sufficient to cause a heterochromatin maintenance impairment (Schalch et al., 2009)), LOOP1B/2B has a heterochromatin defect *in vivo*, indicating that the Chp1CD-Nucleosome core interaction, mediated by LOOP1 and LOOP2, has a broad significance in heterochromatin formation that goes beyond the sole recognition of the H3K9me3 histone tail.

The Nucleosome core interaction positions the Chp1CD in a defined orientation with the positively charged C-terminal α -helix protruding outwards from the Nucleosome surface (data not shown). A previous study (Ishida et al., 2012), showed that Chp1CD is able to interact with nucleic acids. This interaction is mediated by the positively charged lysine residues of the C-terminal α -helix as shown by NMR spectroscopy experiments (Ishida et al., 2012) Figure S4 (C-D)). Mutations aimed at swapping C-terminal α -helix residues between Swi6CD, which does not bind nucleic acids *in vitro*, and Chp1CD (α mut1 and α mut2 mutants) caused not only loss of binding for RNA and DNA *in vitro* but also a severe heterochromatin maintenance defect

in vivo. Notably, both α mut1 and α mut2 mutants bind the H3K9me3 peptide with the same affinity as the wild-type Chp1CD, since none of the residues previously shown to be involved in the histone tail binding affinity was modified. The heterochromatin defect therefore originated by the incapacity of the α mut1 and α mut2 Chp1CD to bind the RNA, as shown also in the *in vivo* centromeric RNA-IP assays (Ishida et al., 2012). The Chp1CD W49Y50/DPS, which substituted the DPS residues found in Swi6 as alignment homologous, with the W49 and Y50 residues, showed a defective centromeric RNA binding *in vitro*, immediately rescued by the addition of H3K9me3 peptide in solution (Ishida et al., 2012). The authors concluded that this differential behavior in the presence or absence of H3K9me histone tail was due to a flexibility change of the W49Y50/DPS when free in solution. When bound to the H3K9me3 peptide the Chp1CD W49Y50/DPS is assuming a conformation that allows it to bind with higher affinity with the RNA. LOOP2B mutant (W49AY50AD51A) has two mutations in common to the mentioned W49Y50/DPS, albeit Alanine residues substitutions. Our peptide binding assays (pull downs and MST, Figure 2.5) indicated that LOOP2B is able to bind the H3K9me3 peptide with comparable affinity to the wild-type protein. Moreover, our RNA EMSA assays indicated that, in the presence of the H3K9me3 peptide in solution, Chp1CD LOOP2B is able to bind to the centromeric RNA with no defects observed (Figure 2.5 E). The Chp1CD LOOP2B mutant was also able to completely rescue the heterochromatin defect in a *chp1Δ* strain, restoring the levels of H3K9me2 at centromeres and completely silencing the centromeric RNAs to wild-type levels. This confirmed the results previously obtained (Ishida et al., 2012). The observation of Chp1CD mutants having impaired silencing at centromeres, despite being able to bind the H3K9 methylated histone tail (LOOP1B/2B and α mut1), indicates that stability of the Chp1CD on the Nucleosome core and centromeric RNA binding might play a common role in heterochromatin formation.

The *in vitro* RNA pulldowns in the presence of H3K9me3 Nucleosomes showed that LOOP1B/2B, defective in Nucleosome core interaction, is not able to strongly bind to the centromeric *dg* RNA as the wild-type protein (Figure 2.14). The same RNA shift assays, in the presence of only the histone H3K9me3 peptide, showed that Chp1CD wild-type and LOOP1B/2B have no significant difference in terms of RNA binding affinity (Figure 2.5 E and Figure 2.13). If no H3K9me3 peptide or Nucleosome is present, LOOP2B mutations alone are sufficient to cause an RNA/DNA binding loss (Ishida et al., 2012). This clearly points toward the importance of the histone H3K9me3 tail to recruit the Chp1CD for Nucleosome core interaction and RNA/DNA binding. Chp1CD can bind to RNA and dsDNA with a similar

affinity in the presence of H3K9me2 peptide ($K_d=0.3 \mu\text{M}$) (Ishida et al., 2012). This points out that Chp1CD might also interact with lesser extent with the Nucleosomal DNA, even though our preliminary structural data and the biochemical data highlight that the main interaction is a direct protein-protein contact with the histone globular domain.

Given the orientation of the Chp1CD in our preliminary cryo-EM maps (data not shown), the C-terminal α -helix, which is facing away from the Nucleosome surface, might indeed form a binding interface for RNA stabilization onto chromatin. The binding surface of Chp1CD on the Nucleosome core is variegated, with one main contact for Chp1CDLOOP1 at the H3 histone globular domain and two minor contacts for Chp1CDLOOP2 and the tip of the C-terminal α -helix, spotted at lower contour level close to histone H2B and H4. Point mutations on the Nucleosome core, histone H3 K79A/T80A/D81A, did not have any effects on the stability of the Chp1CD-H3K9me3 Nucleosome core interaction (Figure 2.4 F). This can be explained from the fact that multiple contact points of Chp1CD with the Nucleosome core were detected (data not shown), and the resolution of our preliminary cryo-EM structure did not allow to pinpoint exactly the histone residues interacting, but only the approximate regions.

A similar result was obtained when mutating the Nucleosome interface of the Sir3BAH domain Nucleosome core complex. An H3K79A mutant did not have any effects on the ability of the Sir3BAH domain to bind the Nucleosome (Wang et al., 2013a). In order to obtain a completely defective Nucleosome core binding, all histone residues interacting should be mutated simultaneously, which could also impair the stability of the Nucleosome particle itself.

Moreover, destabilization of only one of the two interacting Chp1CD LOOPS resulted either in a partial or no heterochromatin formation defect. The indication that to have a silencing impaired phenotype we needed to mutate two different LOOPS in Chp1CD argues that the Nucleosome core interaction requires multiple contact points. In this case, improvement in the cryo-EM maps resolution will prove to be fundamental to better understand which residues are involved in the Chp1CD-Nucleosome core interaction, on the side of the Nucleosome. This suggests multiple possibilities regarding the regulation of the stability of Chp1 on the H3K9me chromatin and could highlight the ability of Chp1CD to read multiple histone marks at the same time, adding more complexity on the mechanism of recruitment of the RITS-Argonaute complex to heterochromatin.

The model of interaction between the chromodomain of Chp1 and the Nucleosome core

The structure of the Chp1CD-H3K9me3 Nucleosome complex gave us the initial evidence that chromodomain interaction with the Nucleosome is not only restricted to the H3K9 methylated histone tail but might interest also several parts of the Nucleosome core surface. Questions still remain on how common this binding mechanism is in other chromodomain or chromatin proteins and how regulated it is. The *in vitro* and *in vivo* experiments showed that the Chp1CD-Nucleosome core interaction has a deep impact on heterochromatin stability and it is therefore necessary for Chp1 to exert its function, at least at centromeres. According to the model we propose (Figure 3.1), Chp1 is recruited through H3K9 methylation to chromatin by direct interaction with the histone tail mediated by the aromatic cage (Y22, W44, Y47 residues) of its chromodomain. The high affinity of this interaction, determined not only by the aromatic cage, but also spatially neighboring residues (E23, V24, N59 and F61), is necessary for initial recognition of the H3K9me mark in heterochromatin formation, but not for maintaining the heterochromatic silencing after its establishment (Schalch et al., 2009). The biochemical data suggested that the main recruitment of Chp1CD occurs through the recognition of the H3K9me histone tail, but that the overall stability of the Chp1-Nucleosome complex is dependent on the Nucleosomal core interaction.

The interaction with the Nucleosome core positions the Chp1CD in a defined orientation (data not shown), with the LOOP1/2 in contact with histone H3, H2B and H4 globular domains (mainly with the H3 loop) and the C-terminal lysine rich α -helix protruding outwards. This helix is therefore exposed for interaction with nucleic acids, particularly centromeric RNAs, already shown to be binding Chp1CD helix in *in vivo* RNA-IP and biochemical experiments (Ishida et al., 2012). As already shown by Ishida and colleagues, the RNA binding of Chp1CD is a necessary feature for heterochromatin maintenance, and *in vivo* experiments showed that mutants unable to bind either DNA or RNA *in vitro*, were incapable to maintain centromeric silencing (Ishida et al., 2012). The phenotype we observe is a defective heterochromatin maintenance, which could similarly reflect the Chp1CD inability of properly binding centromeric RNA. *In vitro* binding assays showed that the LOOP1B/2B mutant is less able to interact with H3K9me3 Nucleosomes and efficiently pull down centromeric *dg* RNAs, compared to the wild-type Chp1 (Figure 2.14).

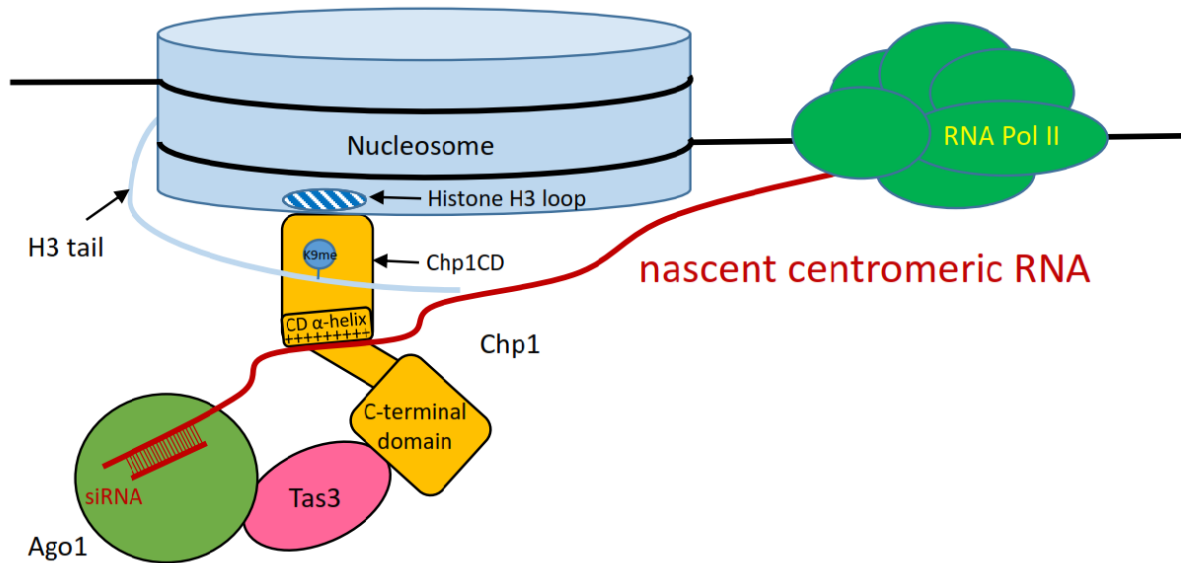


Figure 3.1: Model of interaction of Chp1 - RITS complex with the H3K9 methylated Nucleosome (described in this section)

Therefore, the Nucleosome core binding could determine not only the orientation but also the stability of the Chp1CD for RNA interaction. As shown in Figure 3.1, the binding of RNA on the Chp1CD α -helix gives a stabilized interface for Argonaute siRNA targeting. Chp1 interacts with its C-terminal domain with the GW protein Tas3, forming a platform for Argonaute recruitment (Debeauchamp et al., 2008; Schalch et al., 2011; Verdel et al., 2004). Our model implies that Chp1 stability on the H3K9me3 Nucleosome is dependent on the Nucleosome core interaction. In LOOP1B/2B mutants, where core binding is lost, there might be a defective anchoring of centromeric RNAs onto chromatin, impairing Argonaute siRNA mediated targeting. This would explain why the *in vivo* phenotypes observed previously (Ishida et al., 2012) and in the present study are very similar to each other, with loss of heterochromatin maintenance that might indeed be correlated with a destabilized RNA binding, not related with H3K9me histone tail recruitment (the affinity mutants did not have a phenotype in maintenance (Schalch et al., 2009)). In this case, Argonaute would be less efficiently targeting centromeric RNAs for siRNA amplification and H3K9me mark deposition, which in turn results in H3K9me levels drop (Figure 2.8, LOOP1/2 mutants).

A similar RNA/DNA binding role might be envisioned also for Clr4CD, possessing a similarly lysine-arginine rich C-terminal α -helix and able to bind nucleic acids *in vitro* (Ishida et al., 2012). The chromodomain swapping experiments nonetheless demonstrated that CDs are not equivalent in terms of function: in fact, their primary sequences might have been finely tuned from evolution to adjust their affinities (for both histone tails and the Nucleosomal core) and

coordinate their respective functions at the same heterochromatin loci. Further studies will be needed to see if Nucleosome core binding is a shared feature among fission yeast chromodomain proteins or if it is specific for Chp1.

Conclusions and perspectives

The results of the work presented in this thesis contributed to answer some questions regarding how the affinity of chromodomains for chromatin might be further increased and regulated by their interaction with the Nucleosome core. The advantage of using cryo- electron microscopy to solve the structure of the Chp1CD-H3K9me3 Nucleosome complex is that we would be able to isolate different conformations of binding, as well as different contact points on the Nucleosome surface, allowing us to build a comprehensive model of the interaction, further confirming the biochemical analysis. The model discussed, however, needs to be further implemented and would benefit from the addition of new data regarding the regulation of Chp1CD binding to chromatin. In fact, it is of paramount interest to understand whether post-translational modifications on the so far characterized contact points (data not shown) between Chp1CD and histone globular domains on the Nucleosome core are, in fact, deposited and how they could regulate the binding. This would create a double-switch mechanism of regulation different from the ones previously characterized (Eissenberg, 2012), with the recruiting H3K9me counteracted by a histone globular domain PTM, rather than a modification on a close histone residue (it is the case for double switch modifications such as H3K9me and H3S10 phosphorylation or H3K4me and H3T3 phosphorylation, immediately close to each other). The proposed type of interaction between distant histone marks is similar to what observed for the regulation of the Sir3 protein binding to the Nucleosome in *S. cerevisiae*, with histone H4K16 acetylation and H3K79 methylation acting against Sir3BAH domain Nucleosome core binding and competing with H4K16 deacetylation, operated by Sir2.

Further structural and biochemical studies are needed on Swi6, Chp2 and Clr4 proteins to understand whether the Nucleosome core interaction is specific for Chp1CD or is a conserved feature of chromodomain proteins involved in heterochromatin formation in *S. pombe*. Particularly interesting is the RNA binding role of Clr4CD in heterochromatin formation, not explored in (Ishida et al., 2012). What would be the function of the Clr4 methyltransferase in stabilizing RNA binding onto chromatin? Is Nucleosome core binding a conserved feature between Clr4 and Chp1? How would Nucleosome core binding be affecting the methyltransferase activity of Clr4?

Discussion

Chp1CD binding of RNA has been characterized but the chromodomain is also able to bind the DNA *in vitro* with similar affinity (Ishida et al., 2012). Our study showed that Chp1CD is autonomously able *in vitro* to contact through direct protein-protein interaction the Nucleosome core, even though it is possible that Chp1CD-Nucleosome conformations in which Chp1CD contacts the Nucleosomal DNA exist. There is still therefore the possibility of Chp1CD binding to H3K9me3 tail and contacting not only the transcribed centromeric RNA for Argonaute targeting, but the Nucleosomal DNA of a neighboring Nucleosome. In the case of Clr4CD this would be particularly useful for clustering together the recruiting H3K9me3 Nucleosome and the Nucleosome connected to it, which has to be targeted for H3K9 methylation (Al-Sady et al., 2013). The work presented further discusses how chromodomains evolved in specialized roles and are not interchangeable. RNA/DNA and Nucleosome core binding give additional complexity in our understanding of how chromatin readers' affinity for chromatin is regulated. It has been shown that nucleic acids interaction is not conserved among fission yeast heterochromatin proteins (Ishida et al., 2012). Similarly, Nucleosome core interaction might not be a common feature for heterochromatin proteins in Eukaryotes. Additional structural and biochemical studies are needed to understand the regulation of the diverse functions of chromatin readers in heterochromatin, using both single Nucleosomes and chromatin arrays. These studies will be of fundamental importance to understand how the genetic program of silencing is regulated in different organisms, taking into account the growing complexity, the developmental program, the spatial and time organization of genome expression, in cells and tissues, in health condition and disease.

MATERIALS AND METHODS

MATERIALS

Table 1. Oligonucleotides

DNA oligonucleotides were ordered and purchased from Metabion (Martinsried, Germany).

OLIGONUCLEOTIDES	SEQUENCE
SphI-BamHI chp1+ ORF + endogenous promoter cloning	
346F	CACGCATGCGCAATTATCAAAAGAGGTCGTG
174R	CACGGATCCCTATTTTAAACCAATAGCTCTCATAA AAGT
Chp1 Chromodomain mutants	
349F.R31S	CTGACTCGGTAAATAAAAATGGC
349R(common to R31S, N33A, N35A, LOOP1)	CAAGGATATCTTCCACCTC
351F.N33A	CTGACCGCGTAGCTAAAAATGGC
358F.N35A	CTGACCGCGTAAATAAAGCTGGCATAAAC
359F.E39S	AAACTCGTATTATATAAAATGGGCTGGCTAT
359R.E39S	ATGCCATTTTTATTACGCGGTC
354F.LOOP1A(R31S,N33A,N35A)	CTGACTCGGTAGCTAAAGCTGGCATAAACGAATA T
356F.LOOP2B(W49Y50D51/AA A)	GGCTATGATGCTGCTGCTAATACTTGGGAACCTG AACAAA
356R.LOOP2A/2B	AGCCCATTTTTATATAATATTCGTTTATG
357F.LOOP2A(N52A)	GGCTATGATTGGTATGATGCTACTTGGGAACCT
442F.LOOP1B(N33A,N35A)	CTGACCGCGTAGCTAAAGCTGGCATAAACGAATA T

RT-qPCR and ChIP oligonucleotides	
110A.TDH1_F	CCAAGCCTACCAACTACG
110A.TDH1_R	AGAGACGAGCTTGACGAA
110F.DGF_F	CTGCGGTTCACCCTTAACAT
110F.DGF_R	CAACTGCGGATGGAAAAAGT
110E.DHE_F	GCCCATTCATCAAACGAGTC
110E.DHE_R	GATTCGGCACCTTTGTCATT

Table 2. Plasmids

Plasmid number	Name
p463	pET28a - 6His SUMO Chp1 Chromo-Domain (residues 15-75)
p515	pREP1 - chp1 promoter_chp1_nmt1 terminator
p522	pREP1 - chp1 promoter_chp1 R31S_nmt1 terminator
p523	pREP1 - chp1 promoter_chp1 N33A_nmt1 terminator
p524	pREP1 - chp1 promoter_chp1 E39S_nmt1 terminator
p525	pREP1 - chp1 promoter_chp1 R31SN33AN35A_nmt1 terminator
p526	pREP1 - chp1 promoter_chp1 N52A_nmt1 terminator
p527	pREP1 - chp1 promoter_chp1 W49AY50AD51A_nmt1 terminator
p545	pREP1 - chp1 promoter_chp1 N52A/ R31SN33AN35A_nmt1 terminator

Materials and Methods

p546	pREP1 - chp1 promoter_chp1 W49AY50AD51A/ R31SN33AN35A_nmt1 terminator
p557	pET28a - 6His SUMO Chp1 Chromo-Domain N52A
p558	pET28a - 6His SUMO Chp1 Chromo-Domain W49AY50AD51A
p564	pREP1 - chp1 promoter_chp1 N35A_nmt1 terminator
p569	pET28a - 6His SUMO Chp1 Chromo-Domain W49AY50AD51A/ R31SN33AN35A
p571	pET28a - 6His SUMO Chp1 Chromo-Domain R31SN33AN35A
p576	pET28a - 6His SUMO Chp1 Chromo-Domain R31SN33AN35A/N52A
p597	pET28a - 6His SUMO Chp1 Chromo-Domain R31S
p599	pET28a - 6His SUMO Chp1 Chromo-Domain N35A
p601	pET28a - 6His SUMO Chp1 Chromo-Domain E39S
p606	pET28a - 6His SUMO Chp1 Chromo-Domain N33A
p638	pREP1 - chp1 promoter_chp1 N33AN35A_nmt1 terminator
p639	pREP1 - chp1 promoter_chp1 N33AN35A/W49AY50AD51A_nmt1 terminator
p646	pETDuet - 10His SUMO Chp1 Chromo-Domain N33AN35A
p647	pETDuet - 10His SUMO Chp1 Chromo-Domain N33AN35A/W49AY50AD51A
p406	pET3- <i>Xenopus laevis</i> histone H2A (same as in (Dyer et al., 2004))
p407	pET3- <i>Xenopus laevis</i> histone H2B (same as in (Dyer et al., 2004))
p408	pET3- <i>Xenopus laevis</i> histone H3 (same as in (Dyer et al., 2004))
p409	pET3- <i>Xenopus laevis</i> histone H4 (same as in (Dyer et al., 2004))

p420	pUC57 – 8 X 145bp 601 sequence (plasmid construct received from Davey CA and published in (Vasudevan et al., 2010))
p794	pREP1-(SphI) Region at the 5' of Chp1 gene (Chromosome I, 2215500 – 2215055) (AscI) – HphMX6 resistance cassette –(SphI) Chp1 endogenous promoter (Chromosome I, 2214829 – 2214664) – <i>chp1</i> + ORF– (BamHI) Chp1 terminator (Chromosome I, 2210976 – 2210582) (BamHI).
p795	pREP1-(SphI) Region at the 5' of Chp1 gene (Chromosome I, 2215500 – 2215055) (AscI) – HphMX6 resistance cassette –(SphI) Chp1 endogenous promoter (Chromosome I, 2214829 – 2214664) – <i>chp1LOOP1B/2B</i> ORF– (BamHI) Chp1 terminator (Chromosome I, 2210976 – 2210582) (BamHI).

Table 3. Strains

Strain number	Name
SP28	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216</i>
SP170	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6</i>
SP170 + p515	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1+</i>
SP170 + p522	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1R31S</i>
SP170 + p523	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1N33A</i>
SP170 + p524	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1E39S</i>
SP170 + p525	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1R31SN33AN35A</i>
SP170 + p526	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1N52A</i>
SP170 + p527	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1 W49AY50AD51A</i>
SP170 + p545	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1N52A/ R31SN33AN35A</i>
SP170 + p546	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1 W49AY50AD51A/ R31SN33AN35A</i>
SP170 + p564	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1N35A</i>

Materials and Methods

SP170 + p638	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1 N33AN35A</i>
SP170 + p639	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::TAP-KanMX6 + chp1 N33AN35A/W49AY50AD51A</i>
SP101	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216</i>
SP64	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δclr4::KanMX6</i>
SP967	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 chp1+::HphMX6- chp1+(ChrI- 2214829- 2210582)</i>
SP968	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 chp1+::HphMX6- chp1LOOP1B/2B (ChrI- 2214829- 2210582)</i>
SP970	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::HphMX6- chp1+(ChrI- 2214829- 2210582)</i>
SP972	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 Δchp1::HphMX6- chp1LOOP1B/2B(ChrI- 2214829- 2210582)</i>
SP974	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 chp1+::HphMX6- chp1+(ChrI- 2214829- 2210582) Δclr4::KanMX6</i>
SP975	<i>h+ leu1-32 ura4-D18 imr1R(NCol)::ura4+ oriI ade6-216 chp1+::HphMX6- chp1LOOP1B/2B (ChrI- 2214829- 2210582) Δclr4::KanMX6</i>
Bacterial strains (<i>Escherichia coli</i>)	
X11-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lac^Iq ZΔM15 Tn10 (Ter^r)]</i>
BL21(DE3)pLysS	<i>F⁻ ompT gal dcm lon hsdS_B(r_B⁻ m_B⁻) λ(DE3) pLysS(cm^R)</i>

Table 4. Buffers and solutions

Name	Composition
YES medium (<i>S.pombe</i>)	5g/l Yeast Extract, 30g/l glucose, 0.225 g/l amino acids (leucine, histidine, lysine), 0.225 g/l adenine (+ 20g/l agar for solid plates, +1% FOA for selection under <i>ura4+</i> expression)
EMMc-leu minimal medium (<i>S.pombe</i>)	12.4 g/l EMM without dextrose (Formedium), 0.225 g/l adenine, 0.225 g/l leucine, 0.056 g/l uracil (+ 20g/l agar for solid plates, +1% FOA for selection under <i>ura4+</i> expression)

Materials and Methods

EMMc-ura minimal medium (<i>S.pombe</i>)	12.4 g/l EMM without dextrose (Formedium), 0.225 g/l adenine, 0.225 g/l leucine, (+ 20g/l agar for solid plates, +1% FOA for selection under <i>ura4+</i> expression)
LB medium (<i>E.coli</i>)	10g/l NaCl, 5g/l Yeast Extract, 10g/l Tryptone
Ampicillin	100mg/ml 1000X stock
Kanamycin	50 mg/ml 1000X stock
Chloramphenicol	34mg/ml 1000X stock in 100% Ethanol
IPTG for bacterial induction	1M IPTG stock, used at 0.2-0.5 mM final concentration
RNA purification	
Lysis buffer	300mM NaOAc pH 5.2, 10mM EDTA, 1% SDS
Chromatin immunoprecipitation (ChIP)	
Lysis buffer	50mM Hepes pH 7.5, 1.5M NaOAc, 5mM MgCl ₂ , 1mM EDTA, 1mM EGTA, 0.1% NP-40 and 1X Protease Inhibitors (Complete EDTA- Roche)
Elution buffer	50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS
Protein purifications (<i>E.coli</i> BL21 pLys expression)	
Lysis buffer	20 mM Hepes pH 7.5, 150 mM NaCl, 20 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF
Binding and Wash buffer	20 mM Hepes pH 7.5, 500 mM NaCl, 20 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF (in the Wash buffer 40 mM Imidazole instead)
Elution buffer	20 mM Hepes pH 7.5, 150 mM NaCl, 300 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF
Dyalisis buffer	20 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM DTT
Nucleosome reconstitution (Dyer et al., 2004; Luger et al., 1999)	
Wash buffer	50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM PMSF, 1 mM β -mercaptoethanol
Guanidinium unfolding buffer	6M Guanidinium HCl, 20 mM Tris-HCl pH 7.5, 5 mM DTT
Octamer reconstitution buffer	2 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 5 mM β -mercaptoethanol
RB-high	10 mM Tris-HCl pH 7.5, 250 mM KCl, 1 mM DTT, 1 mM EDTA
RB-low	10 mM Tris-HCl pH 7.5, 2 M KCl, 1 mM DTT, 1 mM EDTA
MLA alkylation	
Alkylation buffer	1 M HEPES pH 7.8, 4 M Guanidinium HCL, 10 mM D/L-methionine
Agarose gel Electrophoresis	
50X TAE	2M Tris acetate pH 8.2-8.4, 50mM EDTA
SDS PAGE	

10X Tris-glycine running buffer	250 mM Tris-Cl, 2.5 M glycine, 1% SDS
Western Blot	
Amido-black solution	7.5% Acetic acid, 20% Ethanol, 0.1% Amido black
De-staining solution	50% Ethanol, 5% Acetic acid
Blotting buffer	20% Ethanol, 39 mM Glycine, 48 mM Tris base, 0.037% SDS
1X TBS-T	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20
Milk solution	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20 + 2.5% (m/v) Skimmed Milk powder
MST (Microscale Thermophoresis)	
Binding buffer	10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM DTT, 0.05% Tween 20 (+10% Glycerol for Nucleosome binding assays)
Nucleosome Binding	
Binding buffer	20 mM Tris-HCl pH 7.5, 75 mM KCl, 0.5 mM DTT (+20 mM Imidazole for Ni-NTA resin Chp1CD-H3K9me3 NCP complex assembly) – used for all Nucleosome complexes
RNA EMSA Binding	
Binding buffer	20 mM Tris-HCl pH 7.5, 100 mM KCl, 0.5 mM DTT and 3% Glycerol

Table 5. Antibodies

Name	Producer	Source	Dilution
Chromatin immunoprecipitation (ChIP)			
Anti-H3K9me2 (ab1220)	Abcam		0.8 µg per 20 µg of Dynabeads Protein A coupled (Invitrogen)
Anti- Chp1 (ab18191)	Abcam	Rabbit polyclonal	0.8 µg per 20 µg of Dynabeads Protein A coupled (Invitrogen)
Western blot (primary antibodies)			
Anti-H3K9me3 (ab8898)	Abcam	Rabbit polyclonal	1:1000 in 2.5% Skimmed Milk in TBS-T
Anti-H3 (ab12079)	Abcam	Goat polyclonal	1:1000 in 2.5% Skimmed Milk in TBS-T
Anti- Chp1 (ab18191)	Abcam	Rabbit polyclonal	1:1000 in 2.5% Skimmed Milk in TBS-T
Anti-Tubulin (B-7)	Santa Cruz	Mouse monoclonal	1:1000 in 2.5% Skimmed Milk in TBS-T
Western blot (secondary antibodies)			
Rabbit anti-Goat IgG (H+L)-HRP Conjugate (#1721034)	BioRAD	Rabbit	1:3000 in 2.5% Skimmed Milk in TBS-T

Goat anti-Rabbit IgG (H+L)-HRP Conjugate (#1706515)	BioRAD	Goat	1:3000 in 2.5% Skimmed Milk in TBS-T
Goat anti-Mouse IgG (H+L)-HRP Conjugate (#1721011)	BioRAD	Goat	1:3000 in 2.5% Skimmed Milk in TBS-T

METHODS

Strains and plasmids construction

All the plasmids used for the *in vivo* assays in this study were generated by modifying the pREP1 plasmid through insertion of the *chp1+* gene with its endogenous promoter (-949 bp from the start of *chp1+* coding sequence, replacing the original *nmt1+* promoter), using SphI-BamHI restriction sites. For the biochemical analysis of Chp1CD function, Chp1CD was expressed from the pET28a expression vector as previously described (Schalch et al., 2009). Chp1CD mutants for the *in vivo* and biochemical experiments were then generated through inverse PCR (iPCR), using the primers listed in Table 1. The PCR products obtained were then purified and treated with T4 PNK (Roche) for 1h at room temperature, ligated with T4 DNA Ligase (NEB) for 1h at 22°C, and subsequently treated with the DpnI endonuclease at 37°C for 1h. The obtained plasmids were then transformed with the standard heat-shock protocol into *E.coli* X11-Blue for sequencing validation and long-term storage, and in BL21(DE3)pLysS for protein expression and purification.

For the *in vivo* heterochromatin rescue assays, the obtained pREP1- *chp1+* derivative and pREP1- *chp1+* CD mutants were transformed into the SP170 *chp1Δ* strain (Table 3). SP170 cells were grown in YES (25 ml) at OD₆₀₀ 1 at 32 °C. Cells were then harvested and washed in 1.2 M Sorbitol (Roth) twice. Pellets were then re-suspended with 1 ml 1.2 M Sorbitol and 500 ng of plasmid DNA were added and mixed. Transformation was performed through electroporation using the Bio-RAD MicroPulser electroporator (Yeast Sph5 settings) and standard electroporation cuvettes (0.2cm gap- green cap - BioRAD). Cells were then plated and selected on EMMc-leu plates.

For genomic integration, the pREP1 plasmid was modified to replace the *nmt1+* promoter with the following integration cassette: (SphI) Region at the 5' of Chp1 gene (Chromosome I, 2215500 – 2215055) (AscI) – *HphMX6* resistance cassette – (SphI) Chp1 endogenous promoter (Chromosome I, 2214829 – 2214664) – Chp1 coding sequence – (BamHI) Chp1 terminator (Chromosome I, 2210976 – 2210582) (BamHI).

The integration cassette (both the *chp1+* cassette and the *chp1LOOP1B/2B* cassette) was then

PCR amplified and transformed by electroporation, (following the transformation procedure mentioned above), into SP101, SP170 and SP64 strains. Cells were initially plated on YES after transformation and replicated the day after on selective YES+ Hygromycin (50 mg/ml Hygromycin) plates. Single colonies were isolated, PCR screened and sequenced for the genomic insertion of the *HphMX6* resistance cassette and the LOOP1B/2B mutations. All bacterial and yeast strains and the plasmids used in this study are listed in Table 2 and Table 3.

Silencing assays

For silencing assays, cells were grown to OD₆₀₀ 0.7-1 and normalized to a final concentration of 10⁷ cells/ml of culture. For the spotting, 5X 1:10 serial dilutions were made so that the densest spot corresponds to 10⁵ cells and the least dense spot corresponds to 10¹ cells. 3 µl of each dilution was spotted onto non-selective (N/S) (either YES or EMMc-leu), 5-fluorootic acid (FOA) and EMMc-ura plates. Cells were then incubated at 32°C for three to four days and imaged. The SP170 *chp1Δ* strains has a *ura4+* reported gene inserted at the centromeric *imr* repeats. When heterochromatin is defective at centromeres, the *ura4+* gene is expressed and becomes toxic in the presence of 5-FOA. Therefore, cells having defective centromeric heterochromatin assembly are counter-selected on 5-FOA plates. Similarly, expression of the *ura4+* reporter can be detected on EMMc-ura auxotrophic selection.

Total RNA purification and Reverse Transcription (RT)

10 ml of yeast cells were harvested at OD₆₀₀ 0.7-1 and the pellets re-suspended in 500 µl of Lysis buffer (300mM NaOAc pH 5.2, 10mM EDTA, 1% SDS). 500 µl of Phenol-Chloroform-Isoamylalcohol (25:24:1) pH for RNA extraction (Roth) were then added to the re-suspension, mixed and incubated at 65°C for 10 minutes with occasional vortexing. Samples were centrifuged at 15000 rpm in a benchtop centrifuge to separate the aqueous and the organic phase. The aqueous phase was then isolated and cold precipitated by the addition of 2 volumes of 100% Ethanol and 1/10 volume of 3 M NaOAcetate pH 5.2. The RNA pellet was washed once with 75% Ethanol, and then re-suspended in DEPC treated water. RNA concentration was then measured at the Nanodrop, and samples were normalized to 50-100 ng/ µl for DNase treatment. 50 µl were treated with DNase I (Roche) at 37°C for 1h. DNase was then heat inactivated at 75°C for 15 minutes.

100-200 ng of RNA were then used to synthesize cDNA by Reverse Transcription (RT), using

SuperscriptIII (Invitrogen) and 1 pmol of each DNA oligo (specific to the genes of interest- in this case *tdh+*, *dg* and *dh* repeats) and following the manufacturer protocol.

RNA levels were subsequently quantified through quantitative PCR (qPCR – later described), using the Flash Dynamo SyBR Green Master Mix (Byozim) and the *tdh+* gene transcript levels for internal normalization.

Chromatin Immunoprecipitation (ChIP)

10 ml yeast cell cultures at OD₆₀₀ 1.3 were cross-linked in 3% formaldehyde (1.3ml of 37% formaldehyde in 50 ml YES/EMMc-leu) at room temperature for 15 minutes as previously described (Marasovic et al., 2013). Cross-linking was then quenched with 125 mM Glycine (2.5 ml of 2.5M Glycine for 50 ml culture volume) for at least 10 minutes at room temperature. Cells were washed once with water and then harvested for flash freezing in liquid nitrogen or immediate lysis. For lysis, cell pellets were re-suspended in 1.5M Sodium acetate Lysis buffer (50 mM HEPES pH 7.5, 1.5 M Sodium acetate, 5 mM MgCl₂, 2 mM EDTA, 2 mM EGTA, 0.1% NP-40, 20% Glycerol), with the addition of PMSF (Roth) and Complete EDTA Free Protease Inhibitors Cocktail (Roche). After throughout re-suspension, 0.25-0.5 mm diameter Glass beads (Roth) were added to the mixture and cells were lysed with bead-beating (4 cycles, 6.5 m/s, 24X2 rotor) with the MP- Biospec Bead beater (MP-Biomedicals). After mechanical lysis, cells were sonicated on the Bioruptor Standard (Diagenode) with 35 cycles, 30s ON and 30s OFF. Lysates were then spun at 13000 rpm at 4°C for 15 minutes, to get rid of the cell debris and obtain the fragmented chromatin in the supernatant. Samples were normalized with each other according to the total RNA and protein quantity measured on Nanodrop and 50 µl of each normalized lysate were used as INPUT. For immunoprecipitation (IP), lysates were incubated with the antibodies indicated in Table 5, for 2h to over-night at 4°C. Antibodies were incubated, prior to IP, with Protein A Dynabeads (Invitrogen) first for 15 minutes at room temperature and then 1 h at 4°C, with constant rotation (Antibody- Dynabeads ratios are indicated in Table 5). After IP, beads were washed 5X with 1.5M Sodium acetate Lysis buffer and elution was done by adding 150 µl of ChIP Elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) per sample, incubating at 65°C for 15 minutes with mild shaking. Cross-link was reversed (both INPUTs and IP samples) by addition of RNaseA for 30 minutes at 37°C, followed by over-night incubation with ProteinaseK at 65°C. De-crosslinked DNA was then extracted by adding 1 volume of Phenol:Chloroform:IsoamylAlcohol (25:24:1) with subsequent Ethanol precipitation.

DNA levels were then quantified with quantitative PCR (qPCR) and normalized to the *tdh+* constitutive gene DNA levels. Oligonucleotides used for ChIP are listed in Table 1.

Quantitative PCR (qPCR or Real-Time PCR)

qPCR was performed on ChIP samples or after reverse transcription (RT) cDNA samples using the 2X DyNamo Flash SyBR Green Master Mix qPCR kit (BioZym) and the TOptical thermocycler (Biometra). 10 μ l PCR reactions were assembled in a 96-well plate (4titude) using 0.4 μ M forward and reverse primers (listed in Table 1), 4 μ l Template DNA with 1X DyNamo Flash SyBR Green Master Mix. For RT-qPCR, qPCR was performed in duplicate for each cDNA sample, with the RNA before reverse transcription as negative control (to detect for contaminating DNA species, in the absence of reverse transcription). For ChIP assays, duplicates or triplicates were analyzed for both INPUT DNA and IPs. Normalization for both RT and ChIP experiments was done by using primers targeting the *tdh+* constitutive gene, which is unaffected by H3K9 methylation or transcript levels variation in the mutants analyzed. The qPCR protocol applied follows these parameters: an initial step of denaturation at 95°C for 3 minutes, followed by 46X cycles (95°C denaturation 10 seconds, 59°C annealing 20 seconds, 72°C elongation 15 seconds) and final melting temperature calculation step (ramping from 60°C to 95°C), for a total duration of 1h and 16 minutes. Oligonucleotides used for qPCR in this study are listed in Table 1.

Protein purification

Wild type His₆-SUMO-Chp1CD and its mutants were expressed in *E.coli* BL21(DE3)pLysS expression strain from a pET28a expression vector. The construct was designed to introduce a Thrombin cleavage site in between the SUMO tag and the His₆ tag (Schalch et al., 2009). SUMO tag could be removed by using the specific Ulp1 SUMO protease. Cells were induced with the addition of 0.2-0.5mM IPTG and grown at 18°C over-night. After induction, cells were harvested at room temperature, washed, re-suspended in Lysis buffer (20 mM Hepes pH 7.5, 150 mM NaCl, 20 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF) and flash frozen in liquid nitrogen, to induce the lysis. The re-suspension was thawed in ice cold water and incubated for 20-30 minutes at 4°C, to allow the Lysozyme to initially break the bacterial cell wall. Cells were sonicated on the Branson Sonifier with one 5 minutes cycle (Branson Sonifier, output 4,

duty cycle 40) and then harvested by centrifugation at 12000 g for 20 minutes at 4°C. Afterwards, the supernatant was incubated with the Ni-NTA Resin (GE Healthcare), already pre-equilibrated in Binding buffer (20 mM Hepes pH 7.5, 500 mM NaCl, 20 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF), for 30 minutes at 4°C, with end-to-end rotation. The binding occurs through the specific recognition of the NTA conjugated Nickel atoms of the resin by the 6 Histidine residue tag of the His₆-SUMO-Chp1CD expressed proteins. Imidazole competes for the binding by mimicking the Histidine residues (Histidine residues have an Imidazole ring) of the tag and causes the elution of the tagged protein at a concentration higher than 200 mM. After incubation, the Ni-NTA resin was washed 5 times with Binding buffer and once with Wash buffer (20 mM Hepes pH 7.5, 500 mM NaCl, 40 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF). Proteins were eventually eluted with Elution buffer (20 mM Hepes pH 7.5, 500 mM NaCl, 300 mM Imidazole pH 8, 0.5 mM DTT, 1 mM PMSF) and then Dialyzed over-night in a buffer containing 20 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM DTT. Further purification of His₆-SUMO-Chp1CD proteins (both wild-type and mutants) was achieved by gel filtration (Superdex 75 pg; GE Healthcare). For Nucleosome-Peptide binding assays and EM structural studies the Chp1CD was dialyzed in 20 mM HEPES pH 7.5, 75 mM KCl, 0.5 mM DTT. Proteins were either used immediately or flash frozen for long-term storage at -80°C. His⁶-Swi6-FLAG and His⁶-Clr4-FLAG for Nucleosome complex assembly were purified according to the same Chp1CD experimental procedure, using the same buffer conditions. FLAG-Chp2 was over-expressed and purified from *S.pombe* following the conditions described in (Marasovic et al., 2013).

Nucleosome in vitro reconstitution and (H3K₉me₃) methylation

Nucleosome were reconstituted by strictly following the Dyer - Luger protocol (Dyer et al., 2004) with the same *Xenopus laevis* histone expressing constructs as previously described (Luger et al., 1997). For reconstitution, it was used the 601 “Widom/Lowary” sequence as previously reported (Lowary and Widom, 1998; Vasudevan et al., 2010). To generate H3K9 methylated Nucleosomes, we used the protocol devised by Matthew D. Simon (Simon et al., 2007), that employs the chemical alkylation of an H3K9C mutated residue to generate a methyl-lysine analog (MLA analog), having all the biochemical functionality of a H3K9 methylated residue. Different methylation states (me₀, me₁, me₂ and me₃) can be produced by the usage of different alkylating agents. The targeting of the mutated cysteine residue (H3K9C) is highly specific and histone H3 had to be mutated (H3C110A) to remove its natural cysteine and introduce the targeted one.

Histone MLA introduction was checked by mass-spec. The mass-spectrum frequently gives rise to a +42 peak over the expected 15297 Da mass, that does not reflect a true contaminating specie and can be present in the original histone preparation (before alkylation). Once alkylated, H3 histones were lyophilized according to the procedure (Simon et al., 2007) and then used for Nucleosome reconstitution, following the standard Dyer-Luger protocol (Dyer et al., 2004).

Chp1CD -H3K9me3 MLA Nucleosome in vitro complex formation and elution

5 µg of purified Chp1CD were bound to 15 µl of Ni-NTA Resin (GE Healthcare) in a Binding buffer containing 20 mM HEPES pH 7.5, 75 mM KCl, 0.5 mM DTT, 20 mM Imidazole for 30 min at 4 °C. 20 mM Imidazole is the minimal concentration required to avoid non-specific H3K9me3 Nucleosome sticking to the Ni-NTA Resin. After the initial binding, unbound Chp1CD was washed away with 5 volumes of Binding buffer. H3K9me3 MLA (Methyl Lysine Analog) Nucleosomes were then added and incubated in binding buffer for 1h on ice in a final volume of 20 µl, with throughout re-suspension (every 5 minutes by hand ticking). The resin was then centrifuged at slow speed (60g, 10 seconds) and washed for 3 times with at least 5X (v/v) Binding buffer. The SUMO-Chp1CD- H3K9me3 MLA Nucleosome complex was eluted by the addition of Thrombin (Sigma) for 2 hours on ice and the eluate collected after centrifugation through the Micro-BioSpin Columns (Bio-RAD), to separate the Ni-NTA Resin from its flow-through. Fractions were then run on a 15% SDS-Acrylamide gel (SDS-PAGE) to check for the purity and the concentration of the Chp1CD- H3K9me3 MLA Nucleosome complex.

The complex obtained was then used for Negative-stain and Cryo- EM. The same procedure of complex formation was applied also for the biochemical studies involving the Chp1CD mutants, to compare their overall affinity for different Nucleosome species. In addition, the same experimental procedure and same buffer conditions were applied for the assembly of the Swi6-H3K9me3 Nucleosome, Chp2-H3K9me3 Nucleosome and Clr4 -H3K9me3 Nucleosome complexes. The affinity FLAG M2 gel (Sigma) was used for anchoring Swi6, Clr4 and Chp2 for the H3K9me3 Nucleosome pull-down.

Nucleosome Trypsin digestion

Nucleosome with no histone tails were produced by Trypsin digestion by incubation with an immobilized TPCK-Trypsin (Thermoscientific) resin for 2h at room temperature in the standard (20 mM HEPES pH 7.5, 75 mM KCl, 0.5 mM DTT) Binding buffer, with a similar

procedure as reported in (Lee et al., 1993) (Figures 2.4 A and B).

Tryptic digestion of nucleosome was already biochemically and structurally characterized (Iwasaki et al., 2013).

Chp1CD-Nucleosome binding assays and Western blot

To compare the different affinities of the Chp1CD wild-type and mutant proteins for Nucleosomes, binding assays were performed on Ni-NTA Resin as previously described (see the section on Chp1CD - H3K₉me₃ MLA Nucleosome *in vitro* complex formation and elution). Three different Nucleosome species were used: H3K₉me₃ Nucleosomes, unmodified Nucleosomes and Trypsin digested (Tailless) Nucleosomes, to estimate the contribution of Nucleosome core and Histone tails to the Chp1CD interaction. Binding assay were performed in 20 mM HEPES pH 7.5, 100 mM KCl, 0.5 mM DTT, 40 mM imidazole. Resins, Flow-through and INPUTs were then run on SDS-PAGE 15% polyacrylamide gels and then transferred onto a PVDF membrane (Millipore) at 15V constant on a semidry Trans-Blot SD transfer apparatus (Bio-RAD) for 30 minutes in Blotting buffer (20% Ethanol, 39 mM Glycine, 48 mM Tris base, 0.037% SDS). To verify the efficiency of transfer, membranes were stained with the Amido-black solution (7.5% Acetic acid, 20% Ethanol, 0.1% Amido black), following de-staining in 50% Ethanol and 5% Acetic acid. Stained membranes were imaged and used as reference to normalize the Chp1CD protein levels. The membrane was then blocked in 1X TBS-T/ 2.5% Milk solution for 30 minutes at room temperature and incubated O/N with anti-H3 histone (AbCam, 1:1000), or anti-H3K₉me₃ Antibody (AbCam, 1:1000) at 4°C. Followed primary antibody incubation, the membrane was washed 3X 5 minutes in 1XTBS-T and incubated for 1h at room temperature with anti-goat IgG-HRP (BioRad, 1:3000) or anti-rabbit IgG-HRP (BioRad, 1:3000). All antibodies were diluted in 1X TBS-T/ 2.5% Milk.

Protein levels were detected using the ECL Plus Pico solution (Thermoscientific) and images were taken using the LAS-3000 Mini Camera.

All western blots in the study presented were performed following this general procedure. To quantify the expression levels of the Chp1CD mutant protein *in vivo* we run SDS 8% polyacrylamide gels and transferred at 15V constant for 1h. To monitor Chp1 protein levels we used the anti- Chp1 (Abcam) antibody with O/N incubation at 4°C. As loading control, either the Tubulin levels (anti-Tubulin antibody -Santa Cruz) or the Amido-black stained membrane were imaged and quantified.

Chp1CD-H3K9me3 peptide binding assays

1 µg of H3K9me3 biotinylated peptide, the [Lys(Me3)9]-Histone H3 (1-21)-GGK(Biotin) peptide from (Eurogentec), was bound for 1 hour at 4°C to 15 µl of Streptavidin Agarose resin (Thermoscientific), in the regular binding buffer described previously (20 mM HEPES pH 7.5, 75 mM KCl, 0.5 mM DTT). The resin was washed once in 5 volumes of the same buffer to get rid of the unbound peptide and 5 µg of Chp1CD (wild-type and mutant) were then added and incubated for 1 hour on ice under constant mixing. After incubation, the flow-through was collected and the resin washed three times (5X volumes each wash). Resin bound fractions and flow-through were then run on a 15% polyacrylamide SDS gel and gels were Coomassie stained and imaged. Quantification of binding was extrapolated by using the ImageJ software from the Coomassie bands intensities. Binding of each mutant was normalized in relation with its own INPUT and compared for each assay to the wild-type Chp1CD (assuming that the wild-type protein binding efficiency would be 100%).

Microscale Thermophoresis (MST)

In order to avoid interference between the SUMO tag and MST measurements and to better compare the different Chp1CD mutants, the SUMO tag was enzymatically removed by using the Ulp1 protease. Each Chp1CD was treated with Ulp1 O/N at 4°C and the cleaved SUMO tag was separated from the cleaved Chp1CD by using the Ni-NTA resin. 100 µg of Chp1CD (wild-type and the selected mutants) were fluorescently labeled with the MO-L003 Monolith™ Protein Labeling Kit BLUE-NHS (Amine Reactive) following the producer (Nanotemper technologies) procedure. With the use of the Nanodrop1000 software feature “Proteins and Labels” an estimated ratio 1:1 between Chp1CD protein and fluorescent label was estimated. Concentrations of the different labeled Chp1CD proteins were further compared on a 15% polyacrylamide SDS gel to normalize them to 0.1mg/ml. In solution peptide and nucleosome binding assays were assembled in a total volume of 20 µl, with a constant quantity of 300 ng of Chp1CD protein in the MST binding buffer containing 10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM DTT and 0.05% Tween-20 (for nucleosome binding assays 10% final glycerol was supplied). A crescent quantity of [Lys(Me3)9]-Histone H3 (1-21)-GGK(Biotin) peptide (Eurogentec) was added to the Chp1CD binding reaction to create the binding curve for MST measurements: 0 nM, 9 nM, 13 nM, 20 nM, 31 nM, 46 nM, 70 nM, 105 nM, 158 nM, 237 nM, 355 nM, 530 µM, 800 µM, 1.2 µM, 1.8 µM. For H3K9me Nucleosomes the following NCP

dilutions were used: 0 nM, 18 nM, 27 nM, 41 nM, 62 nM, 93 nM, 140 nM, 210 nM, 316 nM, 474 nM, 711 nM, 1.06 μ M, 1.2 μ M, 1.4 μ M, 1.6 μ M, 2.4 μ M. Each measurement was done in five replicates on the NT.115 Monolith instrument (Nanotemper technologies) using standard treated capillaries (Nanotemper Cat#K002). The following settings were applied for each run: 80%LED and 40%MST power, with 30s Laser On time and 5s Laser Off time. Three independent experiments were performed for the wild-type and LOOP1B/2B mutant.

Data were analyzed with the SigmaPlot software version 13.0 (Systat Software, San Jose, CA, USA) and the GraphPad Prism software version 6.00 (GraphPad software, San Diego, CA, USA). Following the observed sigmoidal trend, for the peptide binding assay we fitted the Richard's 5 Parameter Logistic Asymmetric Sigmoidal equation, with the software automatically calculating the dissociation constant (K_d). In the case of the Nucleosome binding assays, where some of the curves were not showing a sigmoidal trend, we fitted the third order polynomial (cubic) equation into the raw data points and extrapolated the dissociation constant by using the 50% binding value on the Y-axis and the "Interpolation" feature of the GraphPad Prism suite.

RNA electrophoretic mobility shift assays (EMSAs)

In vitro transcribed 5' dephosphorylated 30nt and 100nt centromeric RNAs were radiolabeled with 32 P by incubating at 37°C with T4 PNK (Roche) for 1h, using [γ - 32 P] labeled ATP. Reactions were then purified by using Microspin G25 Columns (GE Healthcare), to remove the excess of radioactive isotope. 0.66 pmols of 30 nt radiolabeled *dg* RNA were then incubated with 10 μ M CD-Chp1 (wild-type and mutants) in RNA binding buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 0.5 mM DTT and 3% Glycerol) for 1h on ice. For the 100 nt *dg* RNA binding assay, 2 pmols of radiolabeled RNA were incubated with Chp1CD (wild-type and LOOP1B/2B mutant) in the following ratios: 1:1 (2 pmols Chp1CD), 1:5 (10 pmols Chp1CD), and 1:10 (20 pmols Chp1CD). H3K9me3 peptide (Eurogentec) was added into the binding mix at a 1:1 H3K9me peptide: CD-Chp1 ratio, to mimic the *in vivo* binding interaction. Similarly, H3K9me3 Nucleosomes were added to saturate the CD-Chp1 in a 1:1 ratio in the RNA pull down assays with Nucleosomes. Before performing the EMSA assays, SUMO-Chp1CD protein constructs (wild-type and mutants) were digested over-night at 4 °C with the purified His6 -Ulp1 enzyme in the binding buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 0.5 mM DTT) to remove the SUMO tag. To separate the SUMO tag, the Ulp1 enzyme and the undigested SUMO-Chp1CD, 1:1 v/v Fast-flow Ni-NTA resin (GE Healthcare) was added to

the samples and incubated 2h at 4°C in the previously described binding buffer. The flow-through, containing the Chp1CD was then collected, concentrated and run on a 15% acrylamide gel SDS-PAGE for purity control. Then, either used immediately or flash-frozen for long term storage.

For the *in vitro* RNA pull downs, 1 µg of SUMO-Chp1CD (wild-type and LOOP1B/2B mutant) was bound to 15 µl Ni-NTA resin (GE Healthcare) and the H3K9me3 Nucleosome were added to assemble the Chp1CD-H3K9me3 Nucleosome complex as previously described in this section, under the same buffer conditions (20 mM HEPES pH 7.5, 75 mM KCl, 0.5 mM DTT, 20 mM Imidazole). After 3X washing with 50 µl of binding buffer (to wash away the unbound H3K9me3 Nucleosomes), 2 pmols of ³²P labeled 100nt *dg* RNA were added and incubated with the Chp1CD-Nucleosome resin on ice for 1 hour. The resin was then centrifuged at slow speed (60g, 10 seconds) and the flow-through collected. Resin was then washed for 3 times with at least 3X (v/v) Binding buffer and the Chp1CD-Nucleosome-RNA complex was eluted by addition of 300mM Imidazole buffer (20 mM HEPES pH 7.5, 75 mM KCl, 0.5 mM DTT, 300 mM Imidazole), incubated for 1 hour on ice with 5 minutes interval mixing (bound fraction). In all RNA EMSA experiments, samples were loaded on a 10% TBE Native Acrylamide gel (Bis-Acrylamide ratio 1:29) and run for 2 hours at 10 mA constant at 4°C. After over-night exposition, gels were then scanned using the TyphoonFLA9000 phosphoimager.

Negative stain Electron microscopy (EM)

3 µl of eluted Chp1CD-H3K9me3 Nucleosome complex were applied on a glow-discharged 1 nm Carbon coated copper electron microscopy grid (400 Mesh Cu, Quantifoil). Grids were previously glow-discharged at 100V for 45 seconds. After applying the sample onto the carbon EM grid for 45 seconds, grids were rinsed with water and stained with 2.5% Uranyl Acetate for 15 seconds. Negative stained grids were later screened for image collection on a Morgagni transmission electron microscope (FEI Company). The same procedure was applied for all the Negative stain grids shown in this work of Thesis, regardless of the sample type.

REFERENCES

- Adkins, N.L., and Georgel, P.T. (2011). MeCP2: structure and function. *Biochem. Cell Biol.* 89, 1–11.
- Akhtar, A., Zink, D., and Becker, P.B. (2000). Chromodomains are protein-RNA interaction modules. *Nature* 407, 405–409.
- Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980). The structure of histone H1 and its location in chromatin. *Nature* 288, 675–679.
- Allshire, R.C., Nimmo, E.R., Ekwall, K., Javerzat, J.P., and Cranston, G. (1995). Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev.* 9, 218–233.
- Al-Sady, B., Madhani, H.D., and Narlikar, G.J. (2013). Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Mol. Cell* 51, 80–91.
- Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. *Nature* 502, 462–471.
- Armache, K.-J., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science* 334, 977–982.
- Assland, R., and Stewart, F. (1995). The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res.* 23, 3168–3173.
- Audergon, P.N.C.B., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A.L., and Allshire, R.C. (2015). Restricted epigenetic inheritance of H3K9 methylation. *Science* (80-.). 348, 132–135.
- Ayoub, N., Jeyasekharan, A.D., Bernal, J.A., and Venkitaraman, A.R. (2008). HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature* 453, 682–686.
- Bah, A., Wischnewski, H., Shchepachev, V., and Azzalin, C.M. (2012). The telomeric transcriptome of *Schizosaccharomyces pombe*. *Nucleic Acids Res.* 40, 2995–3005.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.
- Bayne, E.H., White, S.A., Kagansky, A., Bijos, D.A., Sanchez-Pulido, L., Hoe, K.-L., Kim, D.-U., Park, H.-O., Ponting, C.P., Rappsilber, J., et al. (2010). Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. *Cell* 140, 666–677.
- Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P., and Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. *Science* 294, 2539–2542.
- Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol. Cell. Biol.* 26, 2560–2569.
- Bhaumik, S.R., Smith, E., and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat. Struct. Mol. Biol.* 14, 1008–1016.
- Bickmore, W.A., and van Steensel, B. (2013). Genome architecture: domain organization of interphase chromosomes. *Cell* 152, 1270–1284.

References

- Bisht, K.K., Arora, S., Ahmed, S., and Singh, J. (2008). Role of heterochromatin in suppressing subtelomeric recombination in fission yeast. *Yeast* 25, 537–548.
- Blus, B.J., Wiggins, K., and Khorasanizadeh, S. (2011). Epigenetic virtues of chromodomains. *Crit. Rev. Biochem. Mol. Biol.* 46, 507–526.
- Braun, S., Garcia, J.F., Rowley, M., Rougemaille, M., Shankar, S., and Madhani, H.D. (2011). The Cul4-Ddb1(Cdt)² ubiquitin ligase inhibits invasion of a boundary-associated antisilencing factor into heterochromatin. *Cell* 144, 41–54.
- Brogaard, K., Xi, L., Wang, J.-P., and Widom, J. (2012). A map of nucleosome positions in yeast at base-pair resolution. *Nature* 486, 496–501.
- Bühler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* 125, 873–886.
- Buker, S.M., Iida, T., Bühler, M., Villén, J., Gygi, S.P., Nakayama, J.-I., and Moazed, D. (2007). Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. *Nat. Struct. Mol. Biol.* 14, 200–207.
- Buscaino, A., Köcher, T., Kind, J.H., Holz, H., Taipale, M., Wagner, K., Wilm, M., and Akhtar, A. (2003). MOF-regulated acetylation of MSL-3 in the *Drosophila* dosage compensation complex. *Mol. Cell* 11, 1265–1277.
- Buscaino, A., Lejeune, E., Audergon, P., Hamilton, G., Pidoux, A., and Allshire, R.C. (2013). Distinct roles for Sir2 and RNAi in centromeric heterochromatin nucleation, spreading and maintenance. *EMBO J.* 32, 1250–1264.
- Cam, H., and Grewal, S.I.S. (2004). RNA interference and epigenetic control of heterochromatin assembly in fission yeast. *Cold Spring Harb. Symp. Quant. Biol.* 69, 419–427.
- Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I.S. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* 37, 809–819.
- Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol. Cell* 41, 67–81.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039–1043.
- Catania, S., Pidoux, A.L., and Allshire, R.C. (2015). Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. *PLoS Genet.* 11, e1004986.
- Colmenares, S.U., Buker, S.M., Buhler, M., Dlakić, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Mol. Cell* 27, 449–461.
- Cooper, J.P., Nimmo, E.R., Allshire, R.C., and Cech, T.R. (1997). Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* 385, 744–747.
- Costantini, M., Clay, O., Auletta, F., and Bernardi, G. (2006). An isochore map of human chromosomes. *Genome Res.* 16, 536–541.

References

- Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* 20, 1147–1155.
- Dai, J., Sultan, S., Taylor, S.S., and Higgins, J.M.G. (2005). The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev.* 19, 472–488.
- Debeauchamp, J.L., Moses, A., Noffsinger, V.J.P., Ulrich, D.L., Job, G., Kosinski, A.M., and Partridge, J.F. (2008). Chp1-Tas3 interaction is required to recruit RITS to fission yeast centromeres and for maintenance of centromeric heterochromatin. *Mol. Cell. Biol.* 28, 2154–2166.
- Djupeadal, I., Portoso, M., Spåhr, H., Bonilla, C., Gustafsson, C.M., Allshire, R.C., and Ekwall, K. (2005). RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev.* 19, 2301–2306.
- Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T.J. (2003). Chromatin fiber folding: Requirement for the histone H4 N-terminal tail. *J. Mol. Biol.* 327, 85–96.
- Du, J., Zhong, X., Bernatavichute, Y.V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A., et al. (2012). Dual Binding of Chromomethylase Domains to H3K9me2-Containing Nucleosomes Directs DNA Methylation in Plants. *Cell* 151, 167–180.
- Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375, 23–44.
- Eissenberg, J.C. (2012). Structural biology of the chromodomain: form and function. *Gene* 496, 69–78.
- Eissenberg, J.C., and Elgin, S.C.R. (2014). HP1a: a structural chromosomal protein regulating transcription. *Trends Genet.* 30, 103–110.
- Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V., and Elgin, S.C. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 87, 9923–9927.
- Eissenberg, J.C., Morris, G.D., Reuter, G., and Hartnett, T. (1992). The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics* 131, 345–352.
- Ekwall, K., Javerzat, J.P., Lorentz, A., Schmidt, H., Cranston, G., and Allshire, R. (1995). The chromodomain protein Swi6: a key component at fission yeast centromeres. *Science* 269, 1429–1431.
- Ekwall, K., Olsson, T., Turner, B.M., Cranston, G., and Allshire, R.C. (1997). Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* 91, 1021–1032.
- Eltsov, M., Maclellan, K.M., Maeshima, K., Frangakis, A.S., and Dubochet, J. (2008). Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19732–19737.
- Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. *Nature* 421, 448–453.
- Fischer, T., Cui, B., Dhakshnamoorthy, J., Zhou, M., Rubin, C., Zofall, M., Veenstra, T.D., and Grewal, S.I.S. (2009). Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8998–9003.

References

- Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* 17, 1870–1881.
- Flanagan, J.F., Mi, L.-Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 438, 1181–1185.
- Folco, H.D., Pidoux, A.L., Urano, T., and Allshire, R.C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* 319, 94–97.
- Garcia, J.F., Al-Sady, B., and Madhani, H.D. (2015). Intrinsic Toxicity of Unchecked Heterochromatin Spread Is Suppressed by Redundant Chromatin Boundary Functions in *Schizosaccharomyces pombe*. *G3 (Bethesda)*.
- Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. *Nat. Rev. Mol. Cell Biol.* 12, 36–47.
- Gerace, E.L., Halic, M., and Moazed, D. (2010). The methyltransferase activity of Clr4Suv39h triggers RNAi independently of histone H3K9 methylation. *Mol. Cell* 39, 360–372.
- Grewal, S.I., and Klar, A.J. (1996). Chromosomal inheritance of epigenetic states in fission yeast during mitosis and meiosis. *Cell* 86, 95–101.
- Grewal, S.I., and Klar, A.J. (1997). A recombinationally repressed region between *mat2* and *mat3* loci shares homology to centromeric repeats and regulates directionality of mating-type switching in fission yeast. *Genetics* 146, 1221–1238.
- Grewal, S.I., Bonaduce, M.J., and Klar, A.J. (1998). Histone deacetylase homologs regulate epigenetic inheritance of transcriptional silencing and chromosome segregation in fission yeast. *Genetics* 150, 563–576.
- Halic, M., and Moazed, D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell* 140, 504–516.
- Hall, I.M., Shankaranarayana, G.D., Noma, K.-I., Ayoub, N., Cohen, A., and Grewal, S.I.S. (2002). Establishment and maintenance of a heterochromatin domain. *Science* 297, 2232–2237.
- Hatano, A., Matsumoto, M., Higashinakagawa, T., and Nakayama, K.I. (2010). Phosphorylation of the chromodomain changes the binding specificity of Cbx2 for methylated histone H3. *Biochem. Biophys. Res. Commun.* 397, 93–99.
- Hayashi, A., Ishida, M., Kawaguchi, R., Urano, T., Murakami, Y., and Nakayama, J. (2012). Heterochromatin protein 1 homologue Swi6 acts in concert with Ers1 to regulate RNAi-directed heterochromatin assembly. *Proc. Natl. Acad. Sci. U. S. A.* 109, 6159–6164.
- Hiriart, E., Vavasour, A., Touat-Todeschini, L., Yamashita, A., Gilquin, B., Lambert, E., Perot, J., Shichino, Y., Nazaret, N., Boyault, C., et al. (2012). Mmi1 RNA surveillance machinery directs RNAi complex RITS to specific meiotic genes in fission yeast. *EMBO J.* 31, 2296–2308.
- Hirota, T., Lipp, J.J., Toh, B.-H., and Peters, J.-M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* 438, 1176–1180.
- Holmes, A.M., Kaykov, A., and Arcangioli, B. (2005). Molecular and cellular dissection of mating-type switching steps in *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* 25, 303–311.

References

- Holoch, D., and Moazed, D. (2015a). RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* *16*, 71–84.
- Holoch, D., and Moazed, D. (2015b). Small-RNA loading licenses Argonaute for assembly into a transcriptional silencing complex. *Nat. Struct. Mol. Biol.* *22*, 328–335.
- Horita, D.A., Ivanova, A. V., Altieri, A.S., Klar, A.J., and Byrd, R.A. (2001). Solution structure, domain features, and structural implications of mutants of the chromo domain from the fission yeast histone methyltransferase Clr4. *J. Mol. Biol.* *307*, 861–870.
- Hsu, J.Y., Sun, Z.W., Li, X., Reuben, M., Tatchell, K., Bishop, D.K., Grushcow, J.M., Brame, C.J., Caldwell, J.A., Hunt, D.F., et al. (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* *102*, 279–291.
- Hughes, R.M., Wiggins, K.R., Khorasanizadeh, S., and Waters, M.L. (2007). Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 11184–11188.
- Hwang, W.L., Deindl, S., Harada, B.T., and Zhuang, X. (2014). Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature* *512*, 213–217.
- Ishida, M., Shimojo, H., Hayashi, A., Kawaguchi, R., Ohtani, Y., Uegaki, K., Nishimura, Y., and Nakayama, J.-I. (2012). Intrinsic nucleic acid-binding activity of Chp1 chromodomain is required for heterochromatic gene silencing. *Mol. Cell* *47*, 228–241.
- Iwasaki, W., Miya, Y., Horikoshi, N., Osakabe, A., Taguchi, H., Tachiwana, H., Shibata, T., Kagawa, W., and Kurumizaka, H. (2013). Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS Open Bio* *3*, 363–369.
- Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* *295*, 2080–2083.
- Jacobs, S.A., Taverna, S.D., Zhang, Y., Briggs, S.D., Li, J., Eissenberg, J.C., Allis, C.D., and Khorasanizadeh, S. (2001). Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *EMBO J.* *20*, 5232–5241.
- James, T.C., and Elgin, S.C. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol. Cell. Biol.* *6*, 3862–3872.
- Jang, S.M., Azebi, S., Soubigou, G., and Muchardt, C. (2014). DYRK1A phosphorylates histone H3 to differentially regulate the binding of HP1 isoforms and antagonize HP1-mediated transcriptional repression. *EMBO Rep.* *15*, 686–694.
- Jia, S. (2004). RNAi-Independent Heterochromatin Nucleation by the Stress-Activated ATF/CREB Family Proteins. *Science* (80-.). *304*, 1971–1976.
- Jia, S., Yamada, T., and Grewal, S.I.S. (2004). Heterochromatin Regulates Cell Type-Specific Long-Range Chromatin Interactions Essential for Directed Recombination. *Cell* *119*, 469–480.
- Jia, S., Kobayashi, R., and Grewal, S.I.S. (2005). Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. *Nat. Cell Biol.* *7*, 1007–1013.
- Johnson, L., Cao, X., and Jacobsen, S. (2002). Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. *Curr. Biol.* *12*, 1360–1367.

References

- Kan, P.-Y., Lu, X., Hansen, J.C., and Hayes, J.J. (2007). The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. *Mol. Cell. Biol.* *27*, 2084–2091.
- Kanoh, J., and Ishikawa, F. (2001). spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr. Biol.* *11*, 1624–1630.
- Kanoh, J., Sadaie, M., Urano, T., and Ishikawa, F. (2005). Telomere Binding Protein Taz1 Establishes Swi6 Heterochromatin Independently of RNAi at Telomeres. *Curr. Biol.* *15*, 1808–1819.
- Kato, H. (2005). RNA Polymerase II Is Required for RNAi-Dependent Heterochromatin Assembly. *Science* (80-.). *309*, 467–469.
- Kawakami, K., Hayashi, A., Nakayama, J.-I., and Murakami, Y. (2012). A novel RNAi protein, Dsh1, assembles RNAi machinery on chromatin to amplify heterochromatic siRNA. *Genes Dev.* *26*, 1811–1824.
- Kaykov, A., Holmes, A.M., and Arcangioli, B. (2004). Formation, maintenance and consequences of the imprint at the mating-type locus in fission yeast. *EMBO J.* *23*, 930–938.
- Keller, C., Adaixo, R., Stunnenberg, R., Woolcock, K.J., Hiller, S., and Bühler, M. (2012). HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Mol. Cell* *47*, 215–227.
- Kim, D., Blus, B.J., Chandra, V., Huang, P., Rastinejad, F., and Khorasanizadeh, S. (2010). Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nat. Struct. Mol. Biol.* *17*, 1027–1029.
- Kim, H.S., Choi, E.S., Shin, J.A., Jang, Y.K., and Park, S.D. (2004). Regulation of Swi6/HP1-dependent heterochromatin assembly by cooperation of components of the mitogen-activated protein kinase pathway and a histone deacetylase Ctr6. *J. Biol. Chem.* *279*, 42850–42859.
- Klutstein, M., Fennell, A., Fernández-Álvarez, A., and Cooper, J.P. (2015). The telomere bouquet regulates meiotic centromere assembly. *Nat. Cell Biol.* *17*, 458–469.
- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* *184*, 868–871.
- Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* *184*, 865–868.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693–705.
- Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* *410*, 116–120.
- De Lange, T. (2009). How Telomeres Solve the End-Protection Problem. *Science* (80-.). *326*, 948–952.
- Lee, D.Y., Hayes, J.J., Pruss, D., and Wolffe, A.P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* *72*, 73–84.
- Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Curr. Opin. Genet. Dev.* *21*, 175–186.
- Li, C., Mueller, J.E., Elflin, M., and Bryk, M. (2008). Linker histone H1 represses recombination at the ribosomal DNA locus in the budding yeast *Saccharomyces cerevisiae*. *Mol. Microbiol.* *67*, 906–919.

References

- Liu, Y., Lu, C., Yang, Y., Fan, Y., Yang, R., Liu, C.-F., Korolev, N., and Nordenskiöld, L. (2011). Influence of histone tails and H4 tail acetylations on nucleosome-nucleosome interactions. *J. Mol. Biol.* 414, 749–764.
- Lomberk, G., Wallrath, L., and Urrutia, R. (2006a). The Heterochromatin Protein 1 family. *Genome Biol.* 7, 228.
- Lomberk, G., Bensi, D., Fernandez-Zapico, M.E., and Urrutia, R. (2006b). Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat. Cell Biol.* 8, 407–415.
- Lorentz, A., Heim, L., and Schmidt, H. (1992). The switching gene *swi6* affects recombination and gene expression in the mating-type region of *Schizosaccharomyces pombe*. *Mol. Gen. Genet.* 233, 436–442.
- Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42.
- Lu, X., Hamkalo, B., Parseghian, M.H., and Hansen, J.C. (2009). Chromatin condensing functions of the linker histone C-terminal domain are mediated by specific amino acid composition and intrinsic protein disorder. *Biochemistry* 48, 164–172.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.
- Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* 119, 1–16.
- Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: Does the 30-nm fibre exist in vivo? *Curr. Opin. Cell Biol.* 22, 291–297.
- Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). Chromatin as dynamic 10-nm fibers. *Chromosoma* 123, 225–237.
- Maison, C., and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. *Nat. Rev. Mol. Cell Biol.* 5, 296–304.
- Marasovic, M., Zocco, M., and Halic, M. (2013). Argonaute and tricomplex generate dicer-independent priRNAs and mature siRNAs to initiate heterochromatin formation. *Mol. Cell* 52, 173–183.
- Min, J., Zhang, Y., and Xu, R.-M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* 17, 1823–1828.
- Mishima, Y., Watanabe, M., Kawakami, T., Jayasinghe, C.D., Otani, J., Kikugawa, Y., Shirakawa, M., Kimura, H., Nishimura, O., Aimoto, S., et al. (2013). Hinge and chromoshadow of HP1 α participate in recognition of K9 methylated histone H3 in nucleosomes. *J. Mol. Biol.* 425, 54–70.
- Mizuguchi, T., Fudenberg, G., Mehta, S., Belton, J.-M., Taneja, N., Folco, H.D., FitzGerald, P., Dekker, J., Mirny, L., Barrowman, J., et al. (2014). Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* 516, 432–435.
- Mizuguchi, T., Barrowman, J., and Grewal, S.I.S. (2015). Chromosome domain architecture and dynamic organization of the fission yeast genome. *FEBS Lett.*
- Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. *Nature* 457, 413–420.
- Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* 119, 789–802.

References

- Motamedi, M.R., Hong, E.-J.E., Li, X., Gerber, S., Denison, C., Gygi, S., and Moazed, D. (2008). HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. *Mol. Cell* 32, 778–790.
- Muchardt, C., Guilleme, M., Seeler, J.-S., Trouche, D., Dejean, A., and Yaniv, M. (2002). Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1 α . *EMBO Rep.* 3, 975–981.
- Muller, H.J. (1930). Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* 22, 299–334.
- Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* 111, 197–208.
- Munari, F., Rezaei-Ghaleh, N., Xiang, S., Fischle, W., and Zweckstetter, M. (2013). Structural plasticity in human heterochromatin protein 1 β . *PLoS One* 8, e60887.
- Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* 292, 110–113.
- Nielsen, P.R., Nietlispach, D., Mott, H.R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A.G., Murzina, N. V, and Laue, E.D. (2002). Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* 416, 103–107.
- Nimmo, E.R., Pidoux, A.L., Perry, P.E., and Allshire, R.C. (1998). Defective meiosis in telomere-silencing mutants of *Schizosaccharomyces pombe*. *Nature* 392, 825–828.
- Nishibuchi, G., and Nakayama, J. (2014). Biochemical and structural properties of heterochromatin protein 1: understanding its role in chromatin assembly. *J. Biochem.* 156, 11–20.
- Nishibuchi, G., Machida, S., Osakabe, A., Murakoshi, H., Hiragami-Hamada, K., Nakagawa, R., Fischle, W., Nishimura, Y., Kurumizaka, H., Tagami, H., et al. (2014). N-terminal phosphorylation of HP1 α increases its nucleosome-binding specificity. *Nucleic Acids Res.* 42, 12498–12511.
- Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S.I.S. (2004). RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* 36, 1174–1180.
- Noma, K., Cam, H.P., Maraia, R.J., and Grewal, S.I.S. (2006). A role for TFIIC transcription factor complex in genome organization. *Cell* 125, 859–872.
- Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I.S., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* 4, 89–93.
- Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). *Science* 183, 330–332.
- Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. *Nat. Rev. Mol. Cell Biol.* 4, 809–814.
- Oppikofer, M., Kueng, S., and Gasser, S.M. (2013). SIR-nucleosome interactions: structure-function relationships in yeast silent chromatin. *Gene* 527, 10–25.
- Padeken, J., and Heun, P. (2014). Nucleolus and nuclear periphery: velcro for heterochromatin. *Curr. Opin. Cell Biol.* 28, 54–60.

References

- Paro, R., and Hogness, D.S. (1991). The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 88, 263–267.
- Partridge, J.F., Borgström, B., and Allshire, R.C. (2000). Distinct protein interaction domains and protein spreading in a complex centromere. *Genes Dev.* 14, 783–791.
- Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am. J. Hum. Genet.* 31, 106–115.
- Peek, A.S., and Behlke, M.A. (2007). Design of active small interfering RNAs. *Curr. Opin. Mol. Ther.* 9, 110–118.
- Petrie, V.J., Wuitschick, J.D., Givens, C.D., Kosinski, A.M., and Partridge, J.F. (2005). RNA interference (RNAi)-dependent and RNAi-independent association of the Chp1 chromodomain protein with distinct heterochromatic loci in fission yeast. *Mol. Cell. Biol.* 25, 2331–2346.
- Platero, J.S., Hartnett, T., and Eissenberg, J.C. (1995). Functional analysis of the chromo domain of HP1. *EMBO J.* 14, 3977–3986.
- Provost, P., Silverstein, R.A., Dishart, D., Walfridsson, J., Djupedal, I., Kniola, B., Wright, A., Samuelsson, B., Radmark, O., and Ekwall, K. (2002). Dicer is required for chromosome segregation and gene silencing in fission yeast cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16648–16653.
- Ragunathan, K., Jih, G., and Moazed, D. (2014). Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* (80-.). science.1258699 – .
- Rea, S., Eisenhaber, F., O’Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.
- Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. *Science* 297, 1831.
- Robinson, P.J.J., and Rhodes, D. (2006). Structure of the “30 nm” chromatin fibre: a key role for the linker histone. *Curr. Opin. Struct. Biol.* 16, 336–343.
- Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the “30-nm” chromatin fiber: evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6506–6511.
- Rougemaille, M., Shankar, S., Braun, S., Rowley, M., and Madhani, H.D. (2008). Ers1, a rapidly diverging protein essential for RNA interference-dependent heterochromatic silencing in *Schizosaccharomyces pombe*. *J. Biol. Chem.* 283, 25770–25773.
- Rougemaille, M., Braun, S., Coyle, S., Dumesic, P.A., Garcia, J.F., Isaac, R.S., Libri, D., Narlikar, G.J., and Madhani, H.D. (2012). Ers1 links HP1 to RNAi. *Proc. Natl. Acad. Sci. U. S. A.* 109, 11258–11263.
- Sadaie, M., Iida, T., Urano, T., and Nakayama, J.-I. (2004). A chromodomain protein, Chp1, is required for the establishment of heterochromatin in fission yeast. *EMBO J.* 23, 3825–3835.
- Sadaie, M., Kawaguchi, R., Ohtani, Y., Arisaka, F., Tanaka, K., Shirahige, K., and Nakayama, J.-I. (2008). Balance between distinct HP1 family proteins controls heterochromatin assembly in fission yeast. *Mol. Cell. Biol.* 28, 6973–6988.

References

- Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* *436*, 138–141.
- Schalch, T., Job, G., Noffsinger, V.J., Shanker, S., Kuscu, C., Joshua-Tor, L., and Partridge, J.F. (2009). High-Affinity Binding of Chp1 Chromodomain to K9 Methylated Histone H3 Is Required to Establish Centromeric Heterochromatin. *Mol. Cell* *34*, 36–46.
- Schalch, T., Job, G., Shanker, S., Partridge, J.F., and Joshua-Tor, L. (2011). The Chp1–Tas3 core is a multifunctional platform critical for gene silencing by RITS. *Nat. Struct. Mol. Biol.* *18*, 1351–1357.
- Schoeftner, S., and Blasco, M.A. (2009). A “higher order” of telomere regulation: telomere heterochromatin and telomeric RNAs. *EMBO J.* *28*, 2323–2336.
- Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* *18*, 1251–1262.
- Scott, K.C., Merrett, S.L., and Willard, H.F. (2006). A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. *Curr. Biol.* *16*, 119–129.
- Shimada, A., Dohke, K., Sadaie, M., Shinmyozu, K., Nakayama, J.-I., Urano, T., and Murakami, Y. (2009). Phosphorylation of Swi6/HP1 regulates transcriptional gene silencing at heterochromatin. *Genes Dev.* *23*, 18–23.
- Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* *311*, 844–847.
- Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* *128*, 1003–1012.
- Smothers, J.F., and Henikoff, S. (2000). The HP1 chromo shadow domain binds a consensus peptide pentamer. *Curr. Biol.* *10*, 27–30.
- Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., and Li, G. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* *344*, 376–380.
- Steiner, N.C., Hahnenberger, K.M., and Clarke, L. (1993). Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. *Mol. Cell. Biol.* *13*, 4578–4587.
- Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I.S. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 152–157.
- Sugiyama, T., Cam, H.P., Sugiyama, R., Noma, K., Zofall, M., Kobayashi, R., and Grewal, S.I.S. (2007). SHREC, an effector complex for heterochromatic transcriptional silencing. *Cell* *128*, 491–504.
- Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* *9*, 465–476.
- Szerlong, H.J., and Hansen, J.C. (2011). Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. *Biochem. Cell Biol.* *89*, 24–34.

References

- Tadeo, X., Wang, J., Kallgren, S.P., Liu, J., Reddy, B.D., Qiao, F., and Jia, S. (2013). Elimination of shelterin components bypasses RNAi for pericentric heterochromatin assembly. *Genes Dev.* 27, 2489–2499.
- Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414, 277–283.
- Thiru, A., Nietlispach, D., Mott, H.R., Okuwaki, M., Lyon, D., Nielsen, P.R., Hirshberg, M., Verreault, A., Murzina, N. V, and Laue, E.D. (2004). Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J.* 23, 489–499.
- Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* 13, 3822–3831.
- Ushinsky, S.C., Bussey, H., Ahmed, A.A., Wang, Y., Friesen, J., Williams, B.A., and Storms, R.K. (1997). Histone H1 in *Saccharomyces cerevisiae*. *Yeast* 13, 151–161.
- Vasudevan, D., Chua, E.Y.D., and Davey, C.A. (2010). Crystal structures of nucleosome core particles containing the “601” strong positioning sequence. *J. Mol. Biol.* 403, 1–10.
- Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I.S., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 303, 672–676.
- Volpe, T., Schramke, V., Hamilton, G.L., White, S.A., Teng, G., Martienssen, R.A., and Allshire, R.C. (2003). RNA interference is required for normal centromere function in fission yeast. *Chromosome Res.* 11, 137–146.
- Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833–1837.
- Wang, F., Li, G., Altaf, M., Lu, C., Currie, M.A., Johnson, A., and Moazed, D. (2013a). Heterochromatin protein Sir3 induces contacts between the amino terminus of histone H4 and nucleosomal DNA. *Proc. Natl. Acad. Sci. U. S. A.* 110, 8495–8500.
- Wang, J., Tadeo, X., Hou, H., Tu, P.G., Thompson, J., Yates, J.R., and Jia, S. (2013b). Epe1 recruits BET family bromodomain protein Bdf2 to establish heterochromatin boundaries. *Genes Dev.* 27, 1886–1902.
- Wang, J., Lawry, S.T., Cohen, A.L., and Jia, S. (2014). Chromosome boundary elements and regulation of heterochromatin spreading. *Cell. Mol. Life Sci.* 71, 4841–4852.
- Wang, J., Reddy, B.D., and Jia, S. (2015). Rapid epigenetic adaptation to uncontrolled heterochromatin spreading. *Elife* 4.
- Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39, 457–466.
- Whitlock, J.P., and Simpson, R.T. (1976). Removal of histone H1 exposes a fifty base pair DNA segment between nucleosomes. *Biochemistry* 15, 3307–3314.
- Wirén, M., Silverstein, R.A., Sinha, I., Walfridsson, J., Lee, H.-M., Laurenson, P., Pillus, L., Robyr, D., Grunstein, M., and Ekwall, K. (2005). Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. *EMBO J.* 24, 2906–2918.

References

- Wood, V., Gwilliam, R., Rajandream, M.-A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., et al. (2002). The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415, 871–880.
- Xhemalce, B., and Kouzarides, T. (2010). A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Dev.* 24, 647–652.
- Yamada, T., Fischle, W., Sugiyama, T., Allis, C.D., and Grewal, S.I.S. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Mol. Cell* 20, 173–185.
- Yamane, K., Mizuguchi, T., Cui, B., Zofall, M., Noma, K., and Grewal, S.I.S. (2011). Asf1/HIRA facilitate global histone deacetylation and associate with HP1 to promote nucleosome occupancy at heterochromatic loci. *Mol. Cell* 41, 56–66.
- Yap, K.L., Li, S., Muñoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M.J., and Zhou, M.-M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* 38, 662–674.
- Yu, R., Jih, G., Iglesias, N., and Moazed, D. (2014). Determinants of heterochromatic siRNA biogenesis and function. *Mol. Cell* 53, 262–276.
- Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I.S. (2008). Roles of the Ctr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat. Struct. Mol. Biol.* 15, 381–388.
- Zhao, T., and Eissenberg, J.C. (1999). Phosphorylation of heterochromatin protein 1 by casein kinase II is required for efficient heterochromatin binding in *Drosophila*. *J. Biol. Chem.* 274, 15095–15100.
- Zhao, T., Heyduk, T., and Eissenberg, J.C. (2001). Phosphorylation site mutations in heterochromatin protein 1 (HP1) reduce or eliminate silencing activity. *J. Biol. Chem.* 276, 9512–9518.
- Zofall, M., Yamanaka, S., Reyes-Turcu, F.E., Zhang, K., Rubin, C., and Grewal, S.I.S. (2012). RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science* 335, 96–100.

LIST OF ABBREVIATIONS

5-FOA	5-Fluorootic Acid
Å	Angstrom
ACF	ATP utilizing Chromatin remodeling and assembly Factor
ATP	Adenosine triphosphate
Amp	Ampicillin
C-terminal	Carboxyl-terminal
CD	Chromodomain
cenRNA	centromeric RNA
ChIP	Chromatin Immunoprecipitation
Chromo	Chromatin organization modifier
CLRC	Clr4-Rik1-Cul4 complex
CSD	Chromoshadow Domain
DNA	Deoxyribonucleic Acid
DTT	1,4-dithio-D,L-threitol
<i>E.coli</i>	Escherichia coli
EM	Electron Microscopy
EMSA	Electrophoretic Mobility Gel Shift
FSC	Fourier Shell Correlation
H3K9me3	Histone H3 Lysine 9 trimethylation
H3K9 _C me3	Histone H3 Lysine 9 (MLA) trimethylation
His-tag	Histidine tag
HP1	Heterochromatin protein 1
HphMX6	Hygromycin MX6 resistance cassette
IPTG	Isopropyl-b-d-thiogalactoside
Kan	Kanamycin
kDa	kilo-Dalton
lncRNA	long non-coding RNA
MLA	Methyl Lysine Analog
MST	Microscale Thermophoresis
mRNA	messenger RNA
N-terminal	Amino-terminal
NCP	Nucleosome core particle
NCPme	H3K9me3 Nucleosome core particle
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance
OD ₆₀₀	Optical density at 600 nm wavelength
ORF	Open reading frame
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PMSF	Phenylmethanesulfonylfluoride
PolII	RNA polymerase II
PTM	Post-Translational Modification
qPCR	quantitative PCR
RDRC	RNA Dependent RNA polymerase Complex
RITS	RNA-Induced Transcriptional Silencing

RNA	Ribonucleic Acid
RT	Reverse Transcription
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
SDS	Sodium Dodecyl Sulphate
SHREC	Snf2-Histone deacetylase Repressor Complex
SUMO	Small ubiquitin related Modifier
Swi	Switching
tRNA	transfer RNA

LIST OF FIGURES AND TABLES

Figure 1.1: The basic unit of chromatin: the Nucleosome core particle (Page 6)

Figure 1.2: Chromatin comes in two main states: Euchromatin and Heterochromatin (Page 10)

Figure 1.3: Overview of the three main constitutive heterochromatin loci in *S. pombe* (Page 15)

Figure 1.4: The “nascent transcript” positive feedback loop model for heterochromatin formation at centromeres in *S.pombe* (Page 19)

Figure 1.5: HP1 proteins and their domains (Page 25)

Figure 1.6: The chromodomain (CD) specifically recognizes H3K9 methylation (Page 29)

Figure 1.7: Chromodomain proteins in fission yeast (Page 32)

Figure 2.1: *In vitro* assembly of the Chp1CD-H3K9_{me3} Nucleosome complex (Page 37)

Figure 2.2: Sequence alignment of Chp1 chromodomain (Page 41)

Figure 2.3: Mutations in Chp1CD LOOP1 and LOOP2 destabilize the interaction with the Nucleosome (Page 42)

Figure 2.4: Chp1CD - Nucleosome interaction is destabilized in LOOP1/2 mutants (Page 43)

Figure 2.5: LOOP1B/2B mutations do not affect the binding affinity for the H3K9_{me3} peptide or the RNA (Page 45)

Figure 2.6: Chp1CD interaction with the Nucleosome core is required for heterochromatin formation (Page 48)

Figure 2.7: Relative expression of *dg* and *dh* centromeric transcripts in wild-type and Chp1CD mutants (Page 49)

Figure 2.8: Centromeric H3K9me2 levels are decreased in Chp1CD LOOP1/2 mutants (Page 50)

Figure 2.9: Chp1 localization at centromeres is compromised in LOOP1/2 mutants (Page 52)

Figure 2.10: The Chp1 proteins (wild-type and Chp1CD mutants) are expressed at the same levels *in vivo* (Page 53)

Figure 2.11: The genomic integration of LOOP1B/2B mutations causes a heterochromatin defective phenotype (Page 54)

Figure 2.12: Centromeric RNA expression levels in wild-type and Chp1LOOP1B/2B genomic integration strains (Page 55)

Figure 2.13: Chp1CDLOOP1B/2B binds to RNA with a slightly higher affinity than the wild-type Chp1CD (Page 56)

Figure 2.14: RNA pulldowns with Chp1CD (wild-type and LOOP1B/2B mutant) in the presence of H3K9me3 Nucleosomes (Page 57)

Figure 2.15: Over-expression of Chp1, Swi6, Chp2 and Clr4 chromodomain proteins destabilizes centromeric silencing (Page 59)

Figure 2.16: *In vitro* assembly of the Swi6-H3K9me3 Nucleosome complex (Page 62)

Figure 2.17: *In vitro* assembly of the Chp2-H3K9me3 Nucleosome complex (Page 64)

Figure 2.18: *In vitro* assembly of the Clr4-H3K9me3 Nucleosome complex (Page 66)

Figure 2.19: BS³ cross-linking titration experiments for the Swi6-H3K9me3 Nucleosome complex (Page 67)

Figure 2.20: Cross-linking of the Swi6-H3K9me3 Nucleosome complex (Page 68)

Figure 2.21: Negative stain EM of the cross-linked Swi6-H3K9me3 Nucleosome and the Swi6L315D-H3K9me3 Nucleosome complexes (Page 68)

Figure 2.22: *In vitro* reconstitution of H3K9me3 di-Nucleosomes (Page 69)

Figure 2.23: Swi6- H3K9me3 di-Nucleosome complex assembly (Page 70)

Figure 2.24: Negative stain EM of reconstituted H3K9me3 di-Nucleosomes (Page 71)

Figure 3.1: Model of interaction of Chp1-RITS complex with the H3K9 methylated Nucleosome (Page 79)

Table 1. Oligonucleotides (Page 82)

Table 2. Plasmids (Page 83)

Table 3. Strains (Page 85)

Table 4. Buffers and solutions (Page 86)

Table 5. Antibodies (Page 88)

ACKNOWLEDGEMENTS

This journey has been a long one and it meant a lot for me on both the professional and personal level. I will always bring the Munich experience along,

First and foremost, I would like to thank my supervisor Prof. Dr. Mario Halic. He actively offered me help whenever I was needing it and always supported me as a PhD student, coming up with new solutions whenever the projects were stalling and teaching me how to tackle on new challenges. His unbiased approach to data analysis and his ethical scientific attitude gave me an example to look forward to.

I would also like to thank him, for actively being part of the project presented in this Thesis, by processing the cryo-EM data and for always sharing with me ideas and discussing on the direction to follow. Nothing would have been possible without his immense effort. I would like to thank my second Gutachter, Prof. Dr. Roland Beckmann and the other members of my PhD defense committee for giving me scientific advice when needed, contributing in creating the Gene Center's nice collaborative environment and for actively help me in the final stages of my PhD time here in Munich.

I would also like to thank Dr. Otto Berninghausen, for kindly introducing me to the technical aspects of EM, for collecting our cryo-EM data and for the nice time spent together at the EM facility.

I would like to thank the Halic laboratory colleagues, Conny, Luca, Ilaria, Sigrun, Silvija and Elias for sharing some easy going conversation in middle of the usual chaotic lab life and being so nice to me: I will always bring you with me independently of where I will be.

I would like to thank as well some of the members of the Hopfner, Beckmann, Gaul and Wilson labs: Carina, Nadine, Agata, Andrea, Alessio, Marko, Martin, Christian, Anne, Clara, Andre, Sara, Daniel, you guys rock!

My PhD would have not been the same if not for five people. I will take my time on them.

Nives, we shared so much during this 2 years that could not help myself not to call you an amazing friend. You are a fantastic soul, a very good scientist and a wonderful company. Sunset on the Gene Center roof! Nice title for a song.

Paolina, non credevo che avrei mai avuto il sorriso con una pipetta in mano, ma nel luglio del 2013 sei arrivata tu, e ne ho avuto la definitiva conferma... Scherzi a parte, tu mi conosci, non sono serio, ma quando penso a te penso a un' amica per la vita. Ti voglio bene.

Mirellina, you have been a great friend, a PhD "war buddy", when fighting against the never ending frustration of experimental failure. I never met anybody so strong and determined, with

Acknowledgements

the will to keep going forward no matter how hard it is. I truly admire you. We are “Motherlovers for life”.

Alessandro and Caro, you are last but by no means least in my heart. You have been amazing friends and it has been incredible sharing moments with you, here in Munich. For me you are Munich, my first mind association with this wonderful city, it's you guys!. I love you both! Thanks for being such good friends!

There is the people you call family, who may or maybe not blood related, but nonetheless share a huge part of your life.

Matthias, you are the best brother “in soul” one could possibly have. Thanks for being such a great support in these last 4 years. I love you.

Davide, my brother, I will never be grateful enough for the gift of having you in my life. I love you.

Mamma, non ce l'avrei fatta senza il tuo supporto costante e il tuo amore. Tu hai sempre capito quello che provo, con te non ho mai dovuto parlare. La nostra connessione e' sempre chiara dentro me e non svanira' mai. Anche adesso che ti scrivo, vicino ad un traguardo, penso che la persona che sono diventato ti portera' sempre dentro. Questo PhD e' anche tuo. Ti voglio bene. Papa', sei il mio eroe. Un esempio umano cosi' grande che non puo' essere paragonato. Onesta', integrita', dovere verso la famiglia e il lavoro: sono questi i valori che mi hai trasmesso. Ovunque andro' rimarra sempre la domanda: cosa farebbe mio padre? Ti voglio bene.

To Nastya, my love, I will dedicate my life. This PhD represented our meeting point, and it saw our love growing day by day. Now I look forward to our life together. I love you.